STRUCTURES TRIDIMENSIONNELLES DE PEPTIDES PAR RMN ET MODÉLISATION MOLÉCULAIRE

par

Abdesslem Khiat

thèse présentée au département de chimie en vue
de l'obtention du grade de docteur ès sciences (Ph.D.)

FACULTÉ DES SCIENCES
UNIVERSITÉ DE SHERBROOKE

Sherbrooke, Québec, Canada, Mai 1996

National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file *Votre référence*

Our file *Notre référence*

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-612-15443-2

Canada

SOMMAIRE

Dans le cadre de la relation structure-activité, nous avons déterminé et analysé par RMN et modélisation moléculaire les structures de quelques peptides. Dans un premier temps, nous avons déterminé et analysé les structures tridimensionnelles du h-CGRP 1-37, de son antagoniste le h-CGRP 8-37 et des analogues de ce dernier (Ala^{17}-, Ala^{20}- et Ala^{21} -hCGRP 8-37). Nous avons montré que le h-CGRP est caractérisé par un segment N-terminal rigide, par une hélice dans le segment (Val^{8}-Leu^{16}) et par un tournant γ dans le segment (Ser^{19}-Gly^{21}). Par contre la structure du h-CGRP 8-37 est moins définie, notamment la structure en hélice n'est plus présente. Dans ce cas, nous avons suggéré que la structure en hélice est stabilisée par des liens hydrogène qui interviennent entre le pont disulfure (Cys^{2}-Cys^{7}) et le segment en hélice. Dans le cas des analogues Ala^{17}-, Ala^{20}- et Ala^{21}-hCGRP 8-37, nous avons montré la conservation de la structure tridimensionnelle du segment C-terminal (Asn^{31}-Phe^{37}) dans les trois analogues hCGRP 8-37, Ala^{17} - et Ala^{20} - hCGRP 8-37, d'où l'importance du segment C-terminal pour l'activité antagoniste.

Dans un deuxième temps, nous avons déterminé les déplacements chimiques en fonction de la température des protons amides NH de la motiline 1-12 et de dix de ses analogues $(CH_2NH)^{1-2}$-, $(CH_2NH)^{2-3}$ ... $(CH_2NH)^{10-11}$ -motiline 1-12. Par la suite, les structures tridimensionnelles des trois analogues motiline 1-12, $(CH_2NH)^{1-2}$- et $(CH_2NH)^{4-5}$-motiline 1-12 ont été étudiées. Ces données montrent une structure tridimensionnelle similaire pour les analogues motiline 1-12, $(CH_2NH)^{1-2}$- , $(CH_2NH)^{2-3}$-motiline 1-12 sur le segment C-terminal, d'où son importance dans l'activité biologique.

Dans le cadre de la mécanique moléculaire, nous avons exploré l'espace conformationnel accessible au peptide YSPTSPSY, élucidé sa conformation libre et liée à un fragment d'ADN, son mode d'intercalation et les interactions intermoléculaires mises en jeu, ainsi que les changements conformationnels du fragment $d(GACGTC)_2$ de l'ADN en double hélice.

REMERCIEMENTS

Je tiens à exprimer ma profonde reconnaissance aux personnes et institutions qui ont contribué à la réalisation de ce travail, je veux citer particulièrement :

Monsieur Yvan BOULANGER, Directeur de recherche, pour l'accueil qu'il m'a réservé au sein de son laboratoire, l'amitié qu'il m'a manifestée, l'aide et les conseils qu'il n'a cessé de me prodiguer, je voudrais lui manifester ma profonde gratitude.

Mon amie Cristina Sanda pour son soutien moral.

Monsieur Hugues MÉNARD, Directeur académique, pour son aide financière dans les moments difficiles.

Mes amis Irène et Hubert BEREZOWSKI et Franco CAU pour leur aide et leur soutien moral.

L'Institut National de la Recherche Scientifique pour sa contribution financière.

Messieurs Louis SENÉCAL, Alain LAROCQUE pour leur assistance technique et Serge ST-PIERRE pour la synthèse des produits peptidiques.

TABLE DES MATIÈRES

LISTE DES TABLEAUX

LISTE DES FIGURES

CHAPITRE III

CHAPITRE IV

INTRODUCTION

Les protéines et les peptides sont d'abord des enchaînements d'acides aminés. Ces molécules possèdent une structure dans l'espace, précise et unique, qui leur permet d'assumer leur fonction dans l'organisme (enzyme, hormone, etc.). Pour comprendre cette fonction et son mécanisme d'action, il est donc nécessaire de connaitre la structure tridimensionnelle de la protéine. De nombreuses techniques expérimentales et théoriques aident à résoudre le problème de la détermination de la géométrie moléculaire: la diffraction des rayons X (1, 2), le dichroïsme circulaire (3), la spectroscopie par fluorescence ou ultraviolet (4), la résonance magnétique nucléaire (RMN) (5, 6), les méthodes de calculs d'orbitales moléculaires *ab initio* ou semi-empiriques et les techniques fondées sur les principes et les lois de la mécanique moléculaire.

La résonance magnétique nucléaire couplée à un ou plusieurs algorithmes (7, 8, 9) de mécanique moléculaire est la méthode de choix pour la détermination des structures tridimensionnelles en solution. À partir des distances interprotons (dérivées des mesures NOE) et des angles dièdres (dérivées des constantes de couplage J ^{1}H-^{1}H), les structures sont calculées d'une façon itérative par de nombreuses méthodes dont les plus populaires sont la distance-géométrie et la dynamique moléculaire restreinte. Dans ce cadre, les articles I, II et IV portent sur les structures en solution du peptide alternatif du gène de la calcitonine (hCGRP 1-37), du fragment antagoniste hCGRP 8-37 et ses analogues, ainsi que sur la motiline 1-12 et ses analogues. Une étude de la relation entre la fonction et la structure tridimensionnelle identifie certains éléments structuraux essentiels pour l'activité.

Parmi les méthodes théoriques, les techniques fondées sur la mécanique moléculaire offrent un bon compromis entre fiabilité, rapidité et coût. Il s'agit de méthodes empiriques intégrant des paramètres géométriques et énergétiques en mécanique classique. La validité de ces méthodes est justifiée par leur capacité à reproduire, dans de nombreux cas, les géométries moléculaires et leurs énergies corrrespondantes. Dans ce cadre, et dans la première partie de l'article III, une approche basée sur une recherche conformationnelle aléatoire avec minimisations d'énergies a été adoptée pour générer l'échantillon le plus représentatif sous forme d'une population de conformations du peptide libre YSPTSPSY. La deuxième partie porte sur le peptide lié à un fragment d'ADN, ce système a été traité par un calcul d'arrimage basé sur la méthode de Monte-Carlo, le mode d'intercalation dans l'ADN ainsi que les interactions intermoléculaires mises en jeu sont analysés et discutés.

Dans le chapitre I, nous ferons un rappel des structures primaire, secondaire, tertiaire et quaternaire des peptides et une revue générale des différentes méthodes de détermination des géométries moléculaires, nous présenterons les méthodes théoriques et expérimentales. L'accent sera mis sur la RMN du proton à deux dimensions et les méthodes utilisées en modélisation moléculaire.

CHAPITRE I

REVUE DES MÉTHODES DE DÉTERMINATION DES STRUCTURES TRIDIMENSIONNELLES

1.1 STRUCTURES DES PEPTIDES

1.1.1 Structure primaire

Les peptides partagent une même structure fondamentale: tout peptide est un polymère linéaire, formé par l'union séquentielle de motifs élémentaires différents, les vingt acides aminés. De la même manière, les différents acides aminés partagent la même structure: sur un carbone central (-Cα-) sont liés, toujours à la même position relative, un groupement amine ($-NH_2$), un groupement carboxyle (-COOH), un atome d'hydrogène et une chaîne latérale R variable. Les acides aminés constituant un peptide diffèrent donc uniquement par la nature de la chaîne latérale R. Les acides aminés au sein d'un peptide sont unis par une liaison covalente entre le groupement amine de l'un et carboxyle de l'autre: c'est la liaison peptidique. Chaque acide aminé incorporé dans une chaîne peptidique est appelé résidu. Donc, d'une protéine à l'autre, seuls diffèrent le nombre et l'ordre des acides aminés, que l'on appelle la séquence. Au sein d'un peptide, on numérote conventionnellement les acides aminés en partant du résidu de l'extrémité N-terminale jusqu'à l'extrémité C-terminale. Le nombre de résidus d'acides aminés d'un peptide se situe entre 2 et 50 environ. La séquence de résidus d'acides

aminés est appelée structure primaire, c'est le premier niveau d'organisation d'un peptide.

1.1.2 Structure secondaire

Le deuxième niveau d'organisation des peptides est la structure secondaire qui réfère aux sous-structures qui sont communes à plusieurs peptides. En effet, un examen plus détaillé de la chaîne peptidique montre la répétition de structures régulières telles que l'hélice α, le feuillet β, les tournants β et les tournants γ. Une chaîne peptidique repliée en hélice a une trajectoire hélicoïdale avec un pas de 3,6 résidus par tour. Les liaisons hydrogène qui la stabilisent en sont la principale caractéristique (entre le C=O du résidu i et le NH du résidu i+4), (figure 1). Un feuillet β a une structure presque totalement plane. Des fragments de chaîne peptidique linéaires y sont maintenus les uns à côté des autres par des liaisons hydrogène entre les hydrogènes des amines des résidus d'un fragment et les oxygènes des carbonyles des résidus du fragment adjacent (figure 1). Plusieurs fragments de feuillets β peuvent être juxtaposés simultanément. Les structures en tournant β sont des pseudo-cycles à dix membres, formés grâce à un lien hydrogène entre le C=O du résidu i et le NH du résidu i+3 (figure 1). Il existe plusieurs types de tournants β (I, II, III, IV, V, VI, VII, I', II', III'), qui diffèrent seulement par les valeurs des angles de torsion (ϕ, ψ) des résidus i+1et i+2. Les tournants γ sont des pseudo-cycles à sept membres stabilisés par un lien hydrogène entre le C=O du résidu i et le NH du résidu i+2 (figure 1). Ces structures secondaires sont communes aux peptides et aux protéines et peuvent être prédites par certaines méthodes empiriques, celle de Chou-Fasman, par exemple est basée sur l'analyse statistique de la structure cristallographique de 29 protéines contenant 4741 résidus acide aminé.

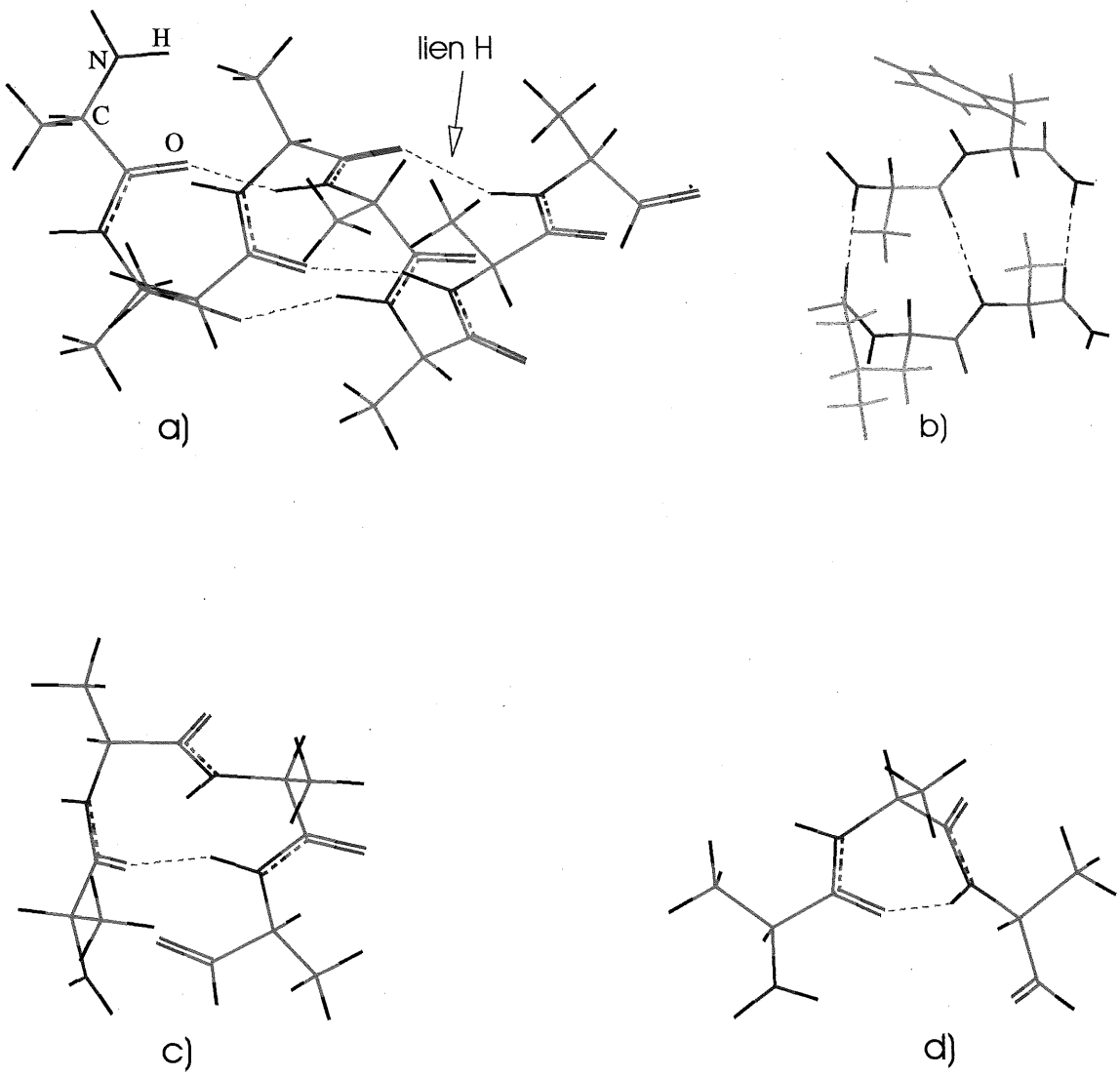


Figure 1 Modèles montrant les structures secondaires

a) Hélice α
b) Feuillet β
c) Tournant β de type I
d) Tournant γ

1.1.3 Structures tertiaire et quaternaire

La structure tridimensionnelle, c'est-à-dire la façon dont s'organisent les éléments de structure secondaire entre eux, est appelée la structure tertiaire. L'existence de dimères ou de multimères de la molécule définit la structure quaternaire.

Les aspects tridimensionnels de la structure des peptides sont issus des possibilités de rotation autour de la plupart des liaisons chimiques de la molécule et sont stabilisés par des liaisons secondaires entre les chaines latérales des résidus : liaisons ioniques, ponts disulfure, ponts hydrogène, interactions hydrophobes. La structure tridimensionnelle d'un peptide peut donc être entièrement définie par sa structure primaire et par l'angle de rotation autour de chacune de ses liaisons. Chaque structure possible pour un peptide est appelée conformation. On peut facilement imaginer un grand nombre de conformations possibles pour chaque peptide, spécialement si le nombre de résidus est élevé. Pourtant, les peptides biologiquement actifs se présentent en général sous une seule ou un nombre limité de conformations. Cette ou ces conformations conditionnent la fonction du peptide dans l'organisme. Tout cela nous permet de poser une question: est-ce que la structure primaire détermine la forme des peptides dans l'espace ? Nous savons, que toute l'information est contenue dans la structure chimique et la séquence primaire, mais nous ne savons pas interpréter cette information. En d'autres termes nous ne connaissons pas les règles qui font qu'un peptide adopte une conformation spatiale particulière parmi les innombrables conformations possibles, en se basant uniquement sur sa séquence d'acides aminés. Heureusement, des techniques expérimentales aussi bien que théoriques ont été développées pour résoudre ce problème fondamental, qui est celui de la détermination de l'organisation spatiale des peptides.

1.2 MÉTHODES DE DÉTERMINATION DES CONFORMATIONS PEPTIDIQUES

1.2.1 Les méthodes expérimentales

1.2.1.1 Le dichroïsme circulaire

La technique du dichroïsme circulaire (3) appliquée aux peptides fournit des renseignements sur le pourcentage de structure secondaire d'un type donné (hélice α, tournant β). Elle ne peut cependant pas localiser ces structures secondaires dans la molécule.

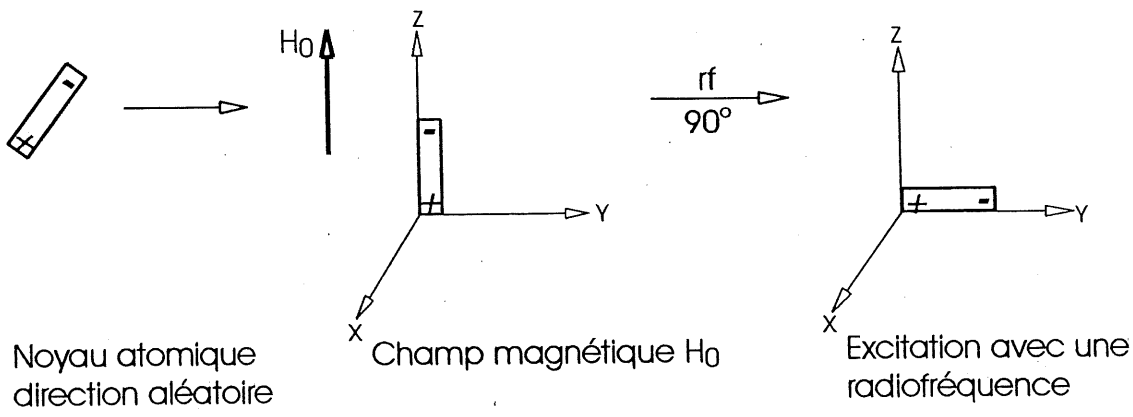
1.2.1.2 La diffraction des rayons X

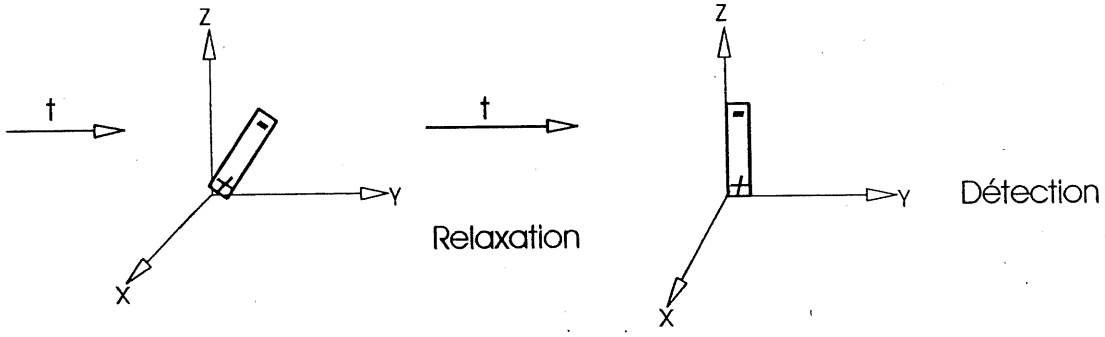
La diffraction des rayons X est la première méthode qui a été utilisée pour la détermination des structures tridimensionnelles des biomolécules (1, 2). Cette méthode est précise mais, malheureusement, de nombreux peptides ne cristallisent pas, ou quand ils cristallisent, la diffraction des monocristaux est insuffisante pour résoudre leur structure. Cette méthode est aussi limitée par le fait qu'elle donne des informations structurales obéissant aux contraintes de l'état solide, qui ne sont pas celles où les peptides existent naturellement.

1.2.1.3 La résonance magnétique nucléaire

Durant les quinze dernières années, le développement de la RMN à deux dimensions (5, 6) a permis un grand progrès dans la détermination des structures tridimensionnelles des biopolymères. La procédure générale est maintenant bien établie (7, 8, 9), 1) identification des systèmes de spin et assignation à un type d'acides aminés, à partir de l'analyse des spectres TOCSY et COSY, 2) détermination des angles dièdres dérivés des constantes de couplage, mesurées à partir du spectre COSY ou unidimensionnel, 3) identification des protons impliqués dans des liaisons hydrogène à partir de la variation des déplacements chimiques des protons NH en fonction de la température, 4) détermination des contraintes de distances interprotons à partir du spectre NOESY qui permettent d'effectuer une assignation séquentielle et 5) utilisation des informations collectées à partir des étapes précédentes comme contraintes pour le calcul des structures qui sont compatibles avec les données RMN.

Si nous plaçons un petit volume de matière contenant des atomes d'hydrogène dans un champ magnétique H_0, le moment magnétique de son noyau (spin) s'aligne dans la direction du champ (le long de l'axe z) (figure 2). En appliquant une onde électromagnétique, le moment magnétique bascule par rapport à H_0. Si le basculement est de 90°, l'impulsion électromagnétique dont la durée provoque un tel basculement est dite impulsion à 90°. Une impulsion à 180° aura une durée double et amènera le moment magnétique dans une direction opposée à sa direction d'équilibre. Quand l'excitation est arrêtée, le moment magnétique revient à sa position d'équilibre en réémettant l'énergie accumulée au cours du basculement, un signal est alors détecté, c'est le signal de précession libre ou FID (*free induction decay*) qui correspond au retour à l'équilibre en fonction du temps. Une transformée de Fourier de ce signal donne le spectre RMN conventionnel ^{1}Hunidimensionnel (1D) (figure2).





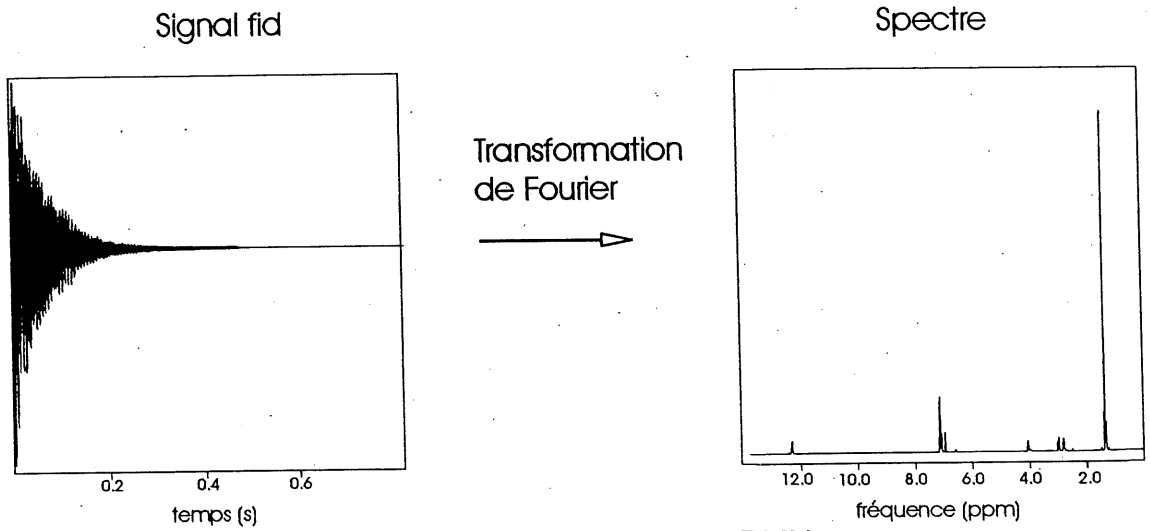


Figure 2 Détection en RMN

Les expériences de RMN à deux dimensions (2D) obéissent aux mêmes schémas de base sauf que les signaux formant la seconde dimension sont affectés par une période de temps variable qui donne lieu à une série de FID. Elles consistent en une période de préparation, une période d'évolution (t_1), une période de mélange et finalement une période de détection (t_2) (figure 3). Une transformée de Fourier dans chaque dimension de $S(t_1, t_2)$, nous donne les spectres à deux dimensions en fréquence $S(\omega_1, \omega_2)$. Dans la plupart des spectres à deux dimensions ^1H-^1H, la diagonale correspond au spectre 1D, les pics placés des deux côtés de la diagonale indiquent l'interaction entre deux protons (figure 3). La nature des interactions dépend du type de l'expérience. Dans le spectre de spectroscopie de corrélation (COSY), les pics correspondent aux couplages entre protons voisins liés chimiquement. À partir de ce type de spectre, on peut aussi mesurer les constantes de couplage pour calculer des angles dièdres (figure 4). Dans la spectroscopie de corrélation totale (TOCSY), les pics correspondent aux connectivités à travers les liaisons chimiques entre protons appartenant à un même système de spin. C'est l'expérience la plus utilisée dans le cas des peptides, le système de spin dans ce cas correspond à un résidu acide aminé (figure 4).

La spectroscopie de l'effet Overhauser nucléaire (NOESY) est la principale source d'information pour déterminer les structures 3D des peptides. Elle nous renseigne sur les protons proches dans l'espace à une distance inférieure ou égale à 5 Å. À cause des mouvements internes de la molécule, cette distance ne peut pas être déterminée précisément, les spectres sont alors interprétés qualitativement. Pour un pic intense, les deux protons couplés sont estimés à une distance inférieure à 3Å, la distance entre les deux protons est de 3 Å à 4 Å pour un pic d'intensité moyenne, et inférieure à 5Å pour une intensité faible. Ce type de spectre nous renseigne aussi sur l'existence des structures secondaires (figure 4). Par exemple, une distance entre les protons amides de la chaîne principale de deux acides aminés i et i+1 (d_{NN}) de 2.8 Å et le(s) proton(s)

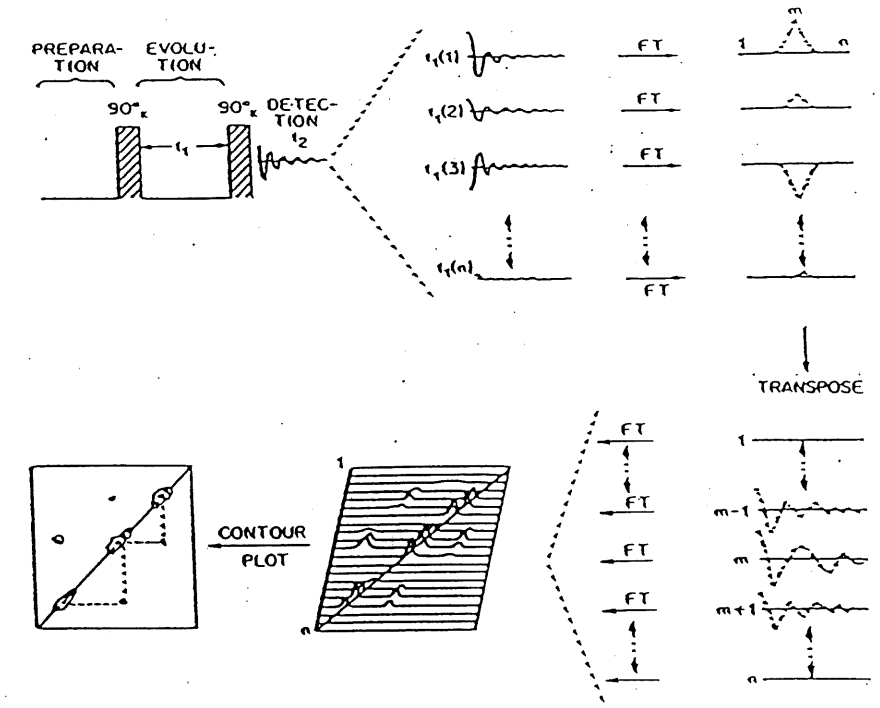
sur le carbone α du résidu i et le(s) proton(s) sur le carbone β du résidu i+1($d_{\alpha\beta}$) de 3.5 Å, qui se répètent sur quatre résidus ou plus nous indique une structure en hélice α (tableau 1).

Tableau 1. Distances ^1H - ^1H dans les structures secondaires des peptides

Distance[a]	Hélice α	Feuillet β	Tournant β I [b]	Tournant β II [b]
$d_{\alpha N}$	3.5	2.2	3.4	2.2
			3.2	3.2
$d_{\alpha N}(i,i+2)$	4.4		3.6	3.3
$d_{\alpha N}(i,i+3)$	3.4		3.1-4.2	3.8-4.7
$d_{\alpha N}(i,i+4)$	4.2			
d_{NN}	2.8	4.3	2.6	4.5
			2.4	2.4
$d_{NN}(i,i+2)$	4.2		3.8	4.3
$d_{\beta N}$	2.5-4.1	3.2-4.5	2.9-4.4	3.6-4.6
			3.6-4.6	3.6-4.6
$d_{\alpha\beta}(i,i+3)$[c]	2.5-4.4			

[a] $d_{AB}(i,j)$: distance du proton A dans le résidu i au proton B dans le résidu j.

[b] Le premier des deux chiffres correspond à la distance entre les résidus 2 et 3 et le deuxième à celle entre les résidus 3 et 4. L'intervalle indiqué pour $d_{\alpha N}(i,i+3)$ correspond aux distances adoptées quand ψ est varié de -180 à 180°. [c] Intervalle de distances adopté par un proton du méthine β si χ_1 est varié de -180 à 180°.



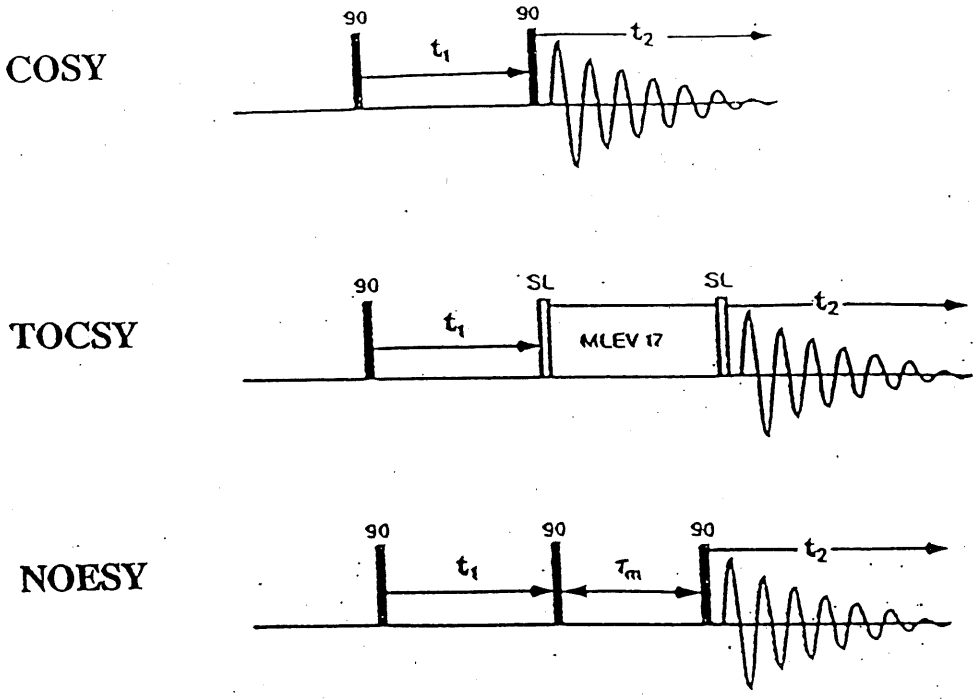
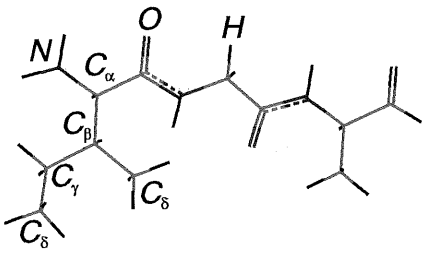
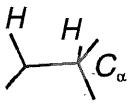
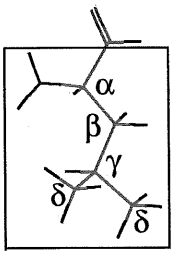
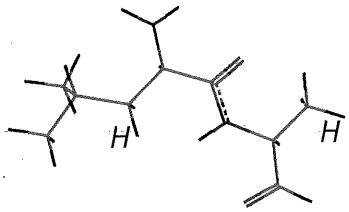


Figure 3 RMN à deux dimensions

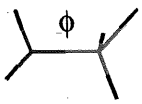


1. Assignation des protons

Technique	Information	
- COSY	Connectivités entre protons voisins couplés	
- TOCSY	Connectivités entre protons dans un même système de spins	
- NOESY	Connectivités entre protons rapprochés ≤ 5 Å	

Résidu n Résidu n+1

2. Information structurale

- NOESY	Protons rapprochés ≤ 5 Å	
- COSY	Constantes de couplage $^3J_{NH\alpha}$ → angles dièdres	

- Coefficients de température → Liaisons hydrogène

Figure 4 Assignation des protons et information structurale

Une autre source d'information peut être obtenue en établissant des spectres 1D à plusieurs températures. Une faible variation négative des déplacements chimiques des protons NH en fonction de la température (supérieure à -3×10^{-3} ppm/K dans le DMSO), nous indique que ces derniers sont en échange lent habituellement parce qu'ils sont impliqués dans des liens hydrogène.

1.2.2 Les méthodes théoriques de modélisation moléculaire

Les méthodes *ab initio*, semi-empiriques et empiriques sont les trois principales méthodes utilisées dans l'analyse conformationnelle des molécules.

1.2.2.1 Les méthodes *ab initio*

Les méthodes *ab initio* consistent à résoudre l'équation de Schrödinger $H\Psi= E\Psi$ et à chercher les orbitales moléculaires qui minimisent l'énergie électronique E. À cause des moyens de calcul insuffisants, ces méthodes restent toujours inadéquates pour l'étude des peptides contenant plusieurs acides aminés.

1.2.2.2 Les méthodes semi-empiriques

Les méthodes semi-empiriques sont basées sur des approximations lors de la résolution de l'équation de Schrödinger. Plusieurs programmes de calcul sont disponibles: MNDO (10), MINDO (11),... Ils sont utilisés pour prédire certaines

propriétés telles que la distribution électronique et les charges partielles. Elles sont aussi utilisées pour minimiser les conformations des molécules, mais sont limitées à l'étude des petites molécules.

1.2.2.3 Les méthodes empiriques

Les méthodes empiriques sont les plus adaptées à l'étude des peptides et des macromolécules en général. Elles utilisent des fonctions énergétiques (champs de force) qui combinent des paramètres déduits de l'expérimentation (rayons X, RMN) et de la théorie (mécanique quantique). Parmi les champs de force disponibles, on peut citer: AMBER (12), CHARMM (13), CVFF (14), ECEPP (15, 16).

- Champ de force ECEPP

Le champ de force ECEPP (*Empirical conformational energy program for peptides*) a été conçu pour l'étude des peptides. Il adopte le modèle à géométrie rigide (les énergies de déformation et d'élongation sont négligées) et l'expression analytique de son énergie totale E_{tot} est la suivante:

$$E_{tot} = 332.0 \sum_i \sum_{j>i} q_i q_j / 2r_{ij} + \sum_i \sum_{j>i} F A^{kl} / r_{ij}^{12} - C^{kl} / r_{ij}^{6} + \sum_i \sum_{j>i} A' / r_{ij}^{12} - B / r_{ij}^{10} + \sum_{\Psi} E_0 (1 + \cos n\Psi) / 2 \quad [1.1]$$

où les différents termes sont, respectivement, l'énergie électrostatique, l'énergie de

van der Waals, l'énergie de liaisons hydrogène et l'énergie de torsion.

- Champ de force CVFF

Le champ de force CVFF (*consistent valence forcefield*) (14) commercialisé par la société BIOSYM est à géométrie flexible, il tient compte des énergies de déformation et d'élongation et son expression analytique est la suivante:

$$E_{tot} = \underset{1}{\sum_b D_b \{1 - e^{-a(b-b0)}\}} + \underset{2}{\sum_\theta H_\theta (\theta-\theta_0)} + \underset{3}{\sum_\phi H_\phi \{1+\cos(n\phi)\}} + \underset{4}{\sum_x H_x x^2}$$

$$[1.2] \quad + \underset{5}{\sum_b \sum_{b'} F_{bb'} (b-b_0)(b'-b'_0)} + \underset{6}{\sum_\theta \sum_{\theta'} F_{\theta\theta'} (\theta-\theta_0)(\theta'-\theta'_0)} + \underset{7}{\sum_b \sum_\theta F_{bq} (b-b_0)(\theta-\theta_0)}$$

$$+ \underset{8}{\sum_\phi F_{\phi\theta\theta'}.\cos\phi\ (\theta-\theta_0)(\theta'-\theta'_0)} + \underset{9}{\sum_x \sum_{x'} F_{xx'} xx'} + \underset{10}{\sum_i \sum_{j>i} \varepsilon \{(r^*/r_{ij})^{12} - 2(r^*/r_{ij})^6\}} + \underset{11}{\sum_i \sum_{j>i} q_i q_j / er_{ij}}$$

Les termes 1-4 sont, respectivement, les énergies de déformation des liaisons, des angles de liaison, des angles de torsion et des interactions hors du plan.
Les termes 5-9 sont les termes croisés et représentent les couplages entre les déformations des coordonnées internes.
Les termes 10-11 décrivent, respectivement, les interactions de van der Waals et les interactions électrostatiques.

Les études menées par Karplus (17) et Hagler (18) montrent que les deux types de champs de force (à géométrie flexible ou rigide) arrivent aux mêmes conclusions

dans l'analyse conformationnelle des macromolécules.

1.3 LES MÉTHODES DE RECHERCHE CONFORMATIONNELLE

Trois méthodes sont utilisées pour générer les conformations d'une molécule: la recherche systématique, la dynamique moléculaire et les méthodes Monte-Carlo.

1.3.1 Les méthodes de recherche systématique

Les méthodes de recherche systématique comprennent la méthode de balayage systématique et la méthode d'enchaînement. La première consiste à faire varier systématiquement chacun des angles de torsion de la molécule pour générer des conformations. Elle est inadéquate dans le cas des macromolécules. En effet, le nombre de conformations augmente considérablement avec l'augmentation de l'espace des angles de torsions (19). La deuxième (20) construit le peptide résidu (minimisé) après résidu (minimisé), et minimise la structure résultante. Son principal inconvénient est le rejet des conformations stabilisées par des interactions entre résidus éloignés.

1.3.2 La dynamique moléculaire

La dynamique moléculaire (21, 22) consiste à simuler le mouvement d'une molécule en fonction du temps, le mouvement de chaque atome étant gouverné par les lois de

La dynamique moléculaire (21, 22) consiste à simuler le mouvement d'une molécule en fonction du temps, le mouvement de chaque atome étant gouverné par les lois de la mécanique de Newton.

Pour une simulation dynamique, il faut :
- choisir une structure initiale, compatible avec les données RMN, par exemple.
- donner des vitesses initiales aux atomes en fixant une température. Ces deux valeurs sont interdépendantes dans l'équation donnant l'énergie cinétique moyenne du système:

[1.3] $$3NKT = \sum_{i=1}^{N} m_i < v_i >^2$$

N étant le nombre d'atomes, K la constante de Boltzmann, T la température, m la masse de l'atome et $< v_i >^2$ la vitesse quadratique moyenne de l' atome i.
- choisir un champ de force qui permet de calculer la force, et par conséquent l'accélération à laquelle est soumis chaque atome. Pour un mouvement uniformément varié:

[1.4] $$F = - \partial E / \partial r = m\gamma = m\, d^2r / dt^2 = \text{constante}$$

E étant l'énergie, r la position atomique, m la masse atomique, γ l'accélération et t le temps.

À la fin de la simulation, les nouvelles positions atomiques sont déterminées par:

[1.5] $$r(t) = 1/2\,\gamma t^2 + v_0 t + r_0$$

où r_0 est la position initiale de l'atome et t le temps de simulation (10 à 200 ps).

On peut accéder par la suite aux différentes conformations de la molécule.

La dynamique moléculaire est très utilisée aussi bien pour les petites molécules que pour les macromolécules. L'exploration correcte de l'espace conformationnel accessible à une molécule dépend du choix du champ de force, du temps de simulation et de la température (23, 24). Ce dernier paramètre conditionne les barrières énergétiques qui peuvent être franchies par la molécule.

1.3.3 Les méthodes Monte-Carlo

La méthode Monte-Carlo consiste à générer d'une façon aléatoire un nombre très important de conformations selon une distribution de Boltzmann. Les propriétés structurales et thermodynamiques sont calculées comme une moyenne de toutes les conformations. Cette méthode est très appropriée pour l'étude des peptides flexibles, qui nécessitent un échantillonnage important (25). Une variante de cette méthode est la méthode Monte-Carlo avec minimisation d'énergie (26). La deuxième partie de l'article III est basée sur ce procédé.

1.3.4 La méthode du recuit simulé

La technique du recuit simulé (*simulated annealing*) (27) couplée à la méthode de Monte-Carlo ou à la dynamique moléculaire exploite le facteur température pour générer les conformations d'une molécule. La simulation est effectuée à haute température (500-1000K), pour un certain intervalle de temps, suivi d'une réduction progressive de la température jusqu'à 300K (28). Cette procédure permet à la molécule de franchir certaines barrières énergétiques.

1.3.5 La méthode PEPSEA

La méthode PEPSEA (*peptide search*) (29) s'inscrit dans le cadre des méthodes empiriques. Elle consiste à générer aléatoirement un grand nombre de conformations de la molécule étudiée et à minimiser leur énergie pour former l'échantillon le plus représentatif de la molécule sous forme d'une population de conformères. Cette population est par la suite analysée par les techniques statistiques, telles que le regroupement des conformères en familles (*clusters*) en tenant compte du facteur entropique.

1.4 Les méthodes couplées avec la RMN

1.4.1 La distance-géométrie

La méthode de distance-géométrie (30, 31) traduit un ensemble de distances interatomiques (les contraintes dérivées des effets NOE) en structures

tridimensionnelles réalistes pour la molécule. À partir d'un nombre limité et imprécis de distances NOE et des distances covalentes, la méthode procède au calcul d'un ensemble de distances plus précises entre tous les atomes par la méthode de l'inégalité triangulaire (32). Elle transforme ensuite ces distances en coordonnées cartésiennes et procède enfin au raffinement des structures par optimisation, en vue de réduire les violations des contraintes (NOE et covalentes) à un niveau acceptable. Parmi les méthodes d'optimisation utilisées, on peut citer la technique de minimisation d'énergie par la méthode du gradient conjugué ou celle de la dynamique moléculaire restreinte. Le dernier-né des programmes basés sur la méthode de distance-géométrie est le DGII (33).

1.4.2 La dynamique moléculaire restreinte

La dynamique moléculaire restreinte (34, 35) utilise la technique du recuit simulé, cette méthode essaie de trouver un compromis entre les contraintes expérimentales et les données énergétiques. Dans l'expression du champ de force, un terme de potentiel relié aux contraintes NOE, E_{NOE}, est ajouté et imposé à la molécule:

[1.6] $$E_{NOE} = \sum_i A_i (d_i - d_{NOE})^2$$

d_i étant la distance entre la ième paire de protons dans les structures dynamiques, d_{NOE} la distance expérimentale entre la même paire de protons et A_i la constante de force. Cette dernière est choisie tel que E_{NOE} tend toujours vers zéro pendant la simulation, et ce, afin que les structures dynamiques satisfassent aux contraintes NOE.

1.4.3 La minimisation d'énergie

Les deux méthodes de minimisation les plus utilisées dans le domaine des biopolymères sont, la méthode de la descente par étape (*steepest descent*) et la méthode du gradient conjugué. Ce sont des techniques de descente itérative qui utilisent des informations issues de la dérivée première de la surface énergétique. Chaque itération se fait en trois étapes consécutives, le choix de la direction de descente r_i, le choix de l'incrément de descente λ_i , et finalement le calcul des nouvelles coordonnées de la molécule représentées par le vecteur v_i à l'itération i. v_i a la même dimension que le nombre de degrés de liberté de la molécule.

[1.7] $$\vec{v}_i = \vec{v}_{i-1} + \lambda_i \vec{r}_i$$

L'incrément λ_i est augmenté par un facteur multiplicateur à chaque itération, tant qu'il conduit à une diminution de l'énergie. Dans le cas d'une augmentation de l'énergie, l'incrément est diminué par un facteur multiplicateur. La seule différence entre la méthode de la descente par étape et la méthode du gradient conjugué réside dans le choix de la direction de descente. La première méthode est généralement utilisée pour une minimisation plus grossière et la dernière méthode pour le raffinement.

1.4.4 La modélisation par homologie

Les algorithmes de détermination des structures conservées par homologie (36, 37) sont basés sur le fait qu'il existe dans des familles de peptides des domaines où les structures sont presque identiques. Une façon de trouver les structures conservées est de calculer les différences de distances moyennes RMSD (*root-mean square deviation*) entre les atomes de la chaîne principale des segments des peptides à comparer, une valeur faible du RMSD (< 1) indiquant que la structure est conservée. En pratique, on compare deux peptides, segment par segment, et on cherche lesquels donnent un RMSD faible entre les coordonnées atomiques de la chaîne principale. Une matrice des distances interatomiques est construite à partir des coordonnées des carbones α, pour les deux peptides à comparer. Par la suite, des petites régions triangulaires des deux matrices sont comparées, le RMSD des éléments des matrices est calculé et le processus se répète. Les valeurs les plus faibles du RMSD sont retenues avec leurs segments (résidus) respectifs dans les deux peptides.

CHAPITRE II

Structure of Human Calcitonin Gene-Related Peptide (hCGRP) and of its Antagonist hCGRP 8-37 as Determined by NMR and Molecular Modeling

Yvan Boulanger, Abdesslem Khiat, Yanmin Chen, Louis Senécal, Yongxue Tu, Serge St-Pierre and Alain Fournier

INRS-Santé, Université du Québec

Peptide Research, vol. 8, No. 4, 206 (1995)

INTRODUCTION

Le peptide alternatif du gène de la calcitonine humaine (hCGRP 1-37) est un peptide de 37 acides aminés qui a plusieurs effets biologiques dont le plus important est la vasodilatation. Le fragment hCGRP 8-37 est un antagoniste du hCGRP 1-37. Le but du travail présenté dans cet article est d'essayer d'expliquer le comportement biologique de ces deux peptides sur des bases structurales, les structures secondaire et tertiaire de ces peptides ont été déterminées par résonance magnétique nucléaire du proton à deux dimensions et par modélisation moléculaire. L'analyse RMN a permis l'assignation et l'identification de plus de 350 connectivités intra- et interrésidus pour chaque peptide. Les modèles moléculaires calculés par dynamique moléculaire et minimisation d'énergie ont permis de montrer que la structure du hCGRP est caractérisée par un segment N-terminal rigide en raison de la présence d'un pont disulfure suivi par un segment en hélice (Val^{8}-Leu^{16}), un tournant γ (Ser^{19}-Gly^{21}) et quelques liaisons hydrogène localisées. La structure du hCGRP 8-37 est moins définie que celle du hCGRP et la structure en hélice n'est plus présente. Les modèles moléculaires des deux peptides confirment les valeurs des coefficients de température des NH et les déplacements chimiques secondaires des protons α. Les liens hydrogène avec le cycle formé par le pont disulfure apparaissent comme critiques dans la formation de la structure en hélice, ces deux éléments de structure étant essentiels pour l'activité agoniste.

Pour la réalisaton de ce travail, j'ai acquis toutes les bases théoriques relatives à la modélisation moléculaire au département de chimie de l'Université de Sherbrooke, j'ai effectué le travail de recherche et de rédaction de la partie modélisation moléculaire et une partie de la RMN à l'INRS-Santé. Les mesures RMN ont été prises à l'Université de Montréal.

ABSTRACT

The solution structures of human calcitonin gene-related peptide (hCGRP, 37 residues) and of its antagonistic fragment hCGRP 8-37 have been determined by two-dimensional ^{1}H NMR spectroscopy and molecular modeling. Analysis of the double quantum filtered correlation spectroscopy, total correlation spectroscopy and nuclear Overhauser enhancement spectroscopy spectra lead to a complete assignment and to the identification of more than 350 intra- and interresidual connectivities for each peptide. Molecular models were calculated by molecular dynamics and energy minimization using distance constraints. The structure of hCGRP is characterized by a rigid N-terminal disulfide-bonded loop followed by helix segments (Val^{8}-Leu^{16}), a γ-turn (Ser^{19}-Gly^{21}) and several local hydrogen bonded patterns. The structure of hCGRP 8-37 is less defined than the structure of hCGRP and no helix structure is present. Molecular models of both peptides are consistent with the NH temperature coefficients and secondary chemical shifts of the α-protons. Hydrogen bonding with the disulfide-bonded ring appears to be critical for helix formation, both structural elements being essential for agonistic activity.

INTRODUCTION

Calcitonin gene-related peptide (CGRP) is a widely-distributed 37 amino acid peptide whose structure is characterized by an amino terminal disulfide-bonded loop (Cys^2-Cys^7). Two forms of the human peptide exist, α- and β-CGRP, which are expressed by separate genes and differ in residue positions 3, 22 and 25 (1, 23). The most notable biological effect of CGRP is on vasodilation but numerous effects on neuronal tissue, skeletal muscle and the cardiovascular system have been reported (11, 13, 22). The C-terminal fragment 8-37 (5, 6, 10, 15) behaves as a competitive antagonist of certain biological effects of CGRP and has lead to the distinction between the $CGRP_1$ and $CGRP_2$ receptors (7).

Structural information for CGRP has been derived from structure-activity studies, spectroscopic (CD and NMR) and theoretical predictive methods. Structure-activity studies have suggested a dependence of the CGRP activity on the helical character of the peptide (18). The existence of an Asp^{14} residue in chicken CGRP, a better helix promoter than Gly^{14} existing in human CGRP (hCGRP), has been suggested to be responsible for the stronger and longer-lasting hypocalcemic and hypophosphatemic activities of the chicken peptide (18). Similarly, replacement of Gly^{23} by Val^{23}, a better helix promoter, in chicken CGRP leads to a two-fold increase in activity (18). In addition, structure-activity studies have demonstrated the importance of the terminal residues (9, 18, 19, 21, 25). Removal of the N-terminal α-amino group improved significantly both biological activity (three-fold in chicken CGRP) and duration of action (18, 19, 21, 25). The C-terminal Phe^{37} is also very sensitive: the addition of electron-withdrawing substituents in the para position of the aromatic ring has been shown to induce up to a five-fold increase in activity for hCGRP (25). On the other hand, removal of C-terminal Phe^{37}-NH_2 causes a large decrease in affinity of the resulting fragment to $CGRP_1$ receptors (20).

Circular dichroism studies of CGRP in aqueous solvent mixtures with 2,2,2-

trifluoroethanol (TFE) (12, 14, 16) or 1,1,1,3,3,3-hexafluoro-2-propanol (HFP) (17) have demonstrated a net solvent-dependent tendency of the chain to be α-helical. These studies also indicated that the existence of an intact disulfide-bridged loop is critical in stabilizing the α-helical conformation. This was emphasized with the results of the CD study of hCGRP 8-37 which showed a weak helical content (17).

The structure of hCGRP has been investigated by two-dimensional ^{1}H NMR spectroscopy in TFE-water solutions (3, 20). Measurements in TFE-water 1: 1 showed that the CGRP structure comprises an amino-terminal disulfide-bonded loop (residues 2-7) leading into a well-defined α-helix between residues 8 and 18, and followed by a predominantly disordered segment although there are indications of a preference for turn-type conformation between residues 19 and 21 (3). More recently, a comparison between CD and NMR data for CGRP 1-37 and CGRP 1-36 revealed that removal of the C-terminal phenylalanine residue does not cause significant structural modifications, although the biological effects are altered (20).

In this study, we investigated the structures of hCGRP (α-form) and of its antagonistic fragment hCGRP 8-37 in DMSO-d_6 using two-dimensional ^{1}H NMR spectroscopy and molecular modeling. Although no solvent can perfectly mimic the receptor site environment, DMSO-d_6 was chosen due to its moderate polarity which allows a stabilization of the conformation of the peptides and should be comparable to the polarity of the receptor site. The structure of hCGRP was found to comprise helical segments following the disulfide-bonded ring in the N-terminal region. The helical structure is destabilized by removal of the disulfide-bonded ring in hCGRP 8-37. Molecular models consistent with all NOE distance constraints and amide temperature coefficients could be calculated for both hCGRP and hCGRP 8-37.

	1	10	20	30
hCGRP	ACDTATCVT HRLAGLLSRS GGVVKNNFVP TNVGSKAF-NH_2			
hCGRP 8-37	VT HRLAGLLSRS GGVVKNNFVP TNVGSKAF-NH_2			

MATERIALS AND METHODS

Peptide Synthesis

Human α-CGRP (hCGRP) and hCGRP 8-37 were synthesized by the solid-phase peptide synthesis method as described previously (17). Briefly, the synthesis was performed with a homemade manual multireactor synthesizer using *tert*-butyloxycarbonyl (Boc) amino acid derivatives, *p*-methylbenzhydrylamine resin as solid support and benzotriazol-1-yloxy-tris-(dimethylamino)phosphonium hexafluorophosphate (BOP) as coupling agent. The peptides were deprotected and cleaved from the resin using hydrofluoric acid in the presence of m-cresol. The formation of the disulfide bridge of hCGRP was carried out in a mixed-solvent system of degassed 80% acetic acid solution (1 mg/ml peptide) and an equal volume of iodine solution (13 mM). After incubation for 2.5 hours with occasional shaking, zinc dust was added until discoloration of the mixture. Crude peptides were finally purified by preparative C_{18} reverse-phase HPLC. Characterization of the peptides was performed by amino acid analysis, analytical HPLC and capillary electrophoresis. The purity of these peptides was estimated to be up to 98%.

NMR Sample Preparation

Synthetic hCGRP and hCGRP 8-37 (5 mg) were dissolved in 100% DMSO-d_6 (0.5 ml) yielding concentrations of 2.6 and 3.3 mM, respectively. The sample solutions were introduced in 5 mm NMR tubes and purged with argon gas to remove oxygen.

Data Acquisition

Two-dimensional ^{1}H NMR spectra of hCGRP and hCGRP 8-37 were recorded at 600.13 MHz and at 298 K on a Bruker AMX 600 (Laboratoire Régional de RMN à Haut Champ, Université de Montréal, Montreal). DQFCOSY, TOCSY (mixing time 60 ms) and NOESY (mixing times 100-300 ms) spectra were collected with 1024 data points and 256 t_1

increments using 32, 64 and 64 transients, respectively. The temperature dependences of the NH chemical shifts were measured on TOCSY spectra acquired at 300, 305, 310, 315, 320, 325 and 330 K on a Bruker AMX2 500 spectrometer operating at 500.13 MHz.

Data Processing

Data processing was performed on Silicon Graphics Indigo R4000 XZ (50 MHz) workstation using the program FELIX 2.30 (Biosym Technologies, Inc., San Diego CA). Prior to Fourier transformation, free induction decays were multiplied with a 30°-shifted sinebell squared window function to improve resolution. For quantitative measurement of NOESY data, a 90°-shifted squared sinebell window function was applied to avoid distortion of peak intensities. Data were zero-filled to yield a 2048 x 2048 matrix with a digital resolution of 3.0 Hz/pt. When required, a baseline correction using a third-order polynomial was applied. Sequential resonance assignments were performed using conventional methods (27) and the main-chain-directed strategy (8). The linewidths measured were relatively sharp (4-10 Hz) and were unaffected by a 10-fold dilution of the samples, suggesting the absence of peptide aggregation. This possibility was therefore not considered in the analysis.

In order to obtain distance constraints, cross-peak intensities were estimated from the 200-ms NOESY spectra. NOE intensities were classified as strong, medium and weak corresponding to upperbound constraints of 3, 4 and 5 Å, respectively. When equivalent proton nuclei were present on a same group, a distance correction corresponding to the maximal distance between the pseudoatom and the protons was added to the constraint (27).

Molecular Modeling

Molecular modeling calculations were carried out on the Silicon Graphics workstation using softwares purchased from Biosym Technologies, Inc. (INSIGHT II, DISCOVER, version 2.35). The extended structure of each peptide, built using the

consistent valence forcefield, was first submitted to 500 iterations of conjugate gradient energy minimization with a RMS gradient of 0.001 Å followed by molecular dynamics simulation using 5 fs time steps to generate 10^5 conformations sampled at every 100 fs. Cluster analysis of these conformations on the basis of interatomic distances allowed to classify them into seven families of conformations. The lowest energy conformation in each family was then energy minimized for 1000 iterations using the steepest descent algorithm. The NMR constraints (NOE distance ranges) were then introduced in each of the seven structures and the molecules were subjected to 30 ps of simulated annealing at 1000K, followed with four cooling steps: from 1000 to 650K for 10 ps, from 650K to 475K for 4 ps, from 475K to 385K for 4 ps and from 385K to 300K for 4 ps. The peptide conformers were finally subjected to 2000 steps of conjugate gradients energy minimization with a RMS gradient of 0.001 Å. Five structures were generated for each of the seven starting conformations. The quality of the final structures was analyzed on the basis of the number of NOE distance violations and backbone RMSD values.

RESULTS AND DISCUSSION

Spectral Assignment

The assignments of the ^{1}H NMR signals of hCGRP (Table 1) and hCGRP 8-37 are presented in Tables 1 and 2, respectively. Most amino acid spin systems were identified from the connectivity patterns observed in DQFCOSY and TOCSY spectra. Aromatic amino acids (His, Phe) were identified using their β-Ar NOESY connectivities. Sequential assignments were performed mostly on the basis of $d_{\alpha N}$, $d_{\beta N}$ and d_{NN} connectivities and were confirmed using connectivities with side chain protons. This procedure allowed the unambiguous assignment of all signals in spite of the presence of multiple identical residues (4 Ala, 4 Asx, 4 Gly, 4 Thr, 5 Val). The serine and arginine sequential assignments were confirmed by a comparison to the [Ala^{17}]-hCGRP 8-37 spectra (data not shown). Signals of lower intensity due to cis Pro^{29} could be observed in the region of residues 27-31. No signal was observed for the C-terminal amide protons.

The chemical shifts of hCGRP 8-37 are similar but not identical to those of hCGRP as shown in Tables 1 and 2. The main differences are observed for residues 8-12. This suggests that the structures of both peptides are slightly different in their common regions. The chemical shift values obtained in DMSO-d_6 differ more significantly from the chemical shifts reported for the same molecule in TFE: water 1: 1 at pH 3.7, even after taking into account solvent effects on chemical shifts (3). This could be explained by structural differences between the two solvent conditions.

Structural Information

The NOESY connectivities observed for hCGRP could be assigned to 211 intraresidual and 150 interresidual proton-proton interactions. The interresidual interactions comprise 88 sequential and 62 medium range (30 i,i+2, 26 i,i+3 and 6 i,i+4) connectivities. In the case of hCGRP 8-37, 174 intraresidual and 167 interresidual connectivities were identified consisting of 106 sequential, 59 medium range (34 i,i+2, 17 i,i+3 and 8 i,i+4)

and two long range connectivities involving Leu[16], Arg[18] and Val[23] . The interresidual connectivities are summarized in the diagonal plots of Figure 1.

Interresidual connectivities indicative of regular secondary structure are presented in Figure 2. In the case of hCGRP, the presence of consecutive intense d_{NN} connectivities between residues 5 and 16 as well as the presence of $d_{\alpha N}(i,i+3)$, $d_{\alpha\beta}(i,i+3)$ and $d_{\alpha N}(i,i+4)$ connectivities in that region suggest the existence of α-helical structure. In the case of hCGRP 8-37, the absence of strong d_{NN} and medium range connectivities is inconsistent with the presence of helical structure.

Figure 3 presents the secondary chemical shifts for the α-protons of hCGRP and hCGRP 8-37 calculated relative to the published random coil chemical shifts in DMSO-d_6 (4). The possibility of helix formation exists for consecutive residues with negative secondary shifts of -0.1 ppm or more (26). These shifts support a modest degree of helix formation between residues 7 and 20 and between residues 23 and 28 for hCGRP. For hCGRP 8-37, the secondary shifts are consistent with little helix except between residues 23 and 26.

The amide protons with low and intermediate temperature coefficients are illustrated in Figure 2. Amide protons displaying the lowest temperature coefficients are those inside and in the vicinity of the disulfide-bonded loop of hCGRP (Thr[4], Thr[6], Cys[7] and Val[8]). Other slow exchanging protons include those of Asn[31] in both peptides as well Gly[21], Phe[27] and Val[32] in hCGRP and Leu[15] in hCGRP 8-37. Amide protons with intermediate temperature coefficients are present at various positions along the chain, especially for hCGRP (Figure 2). The fact that the chemical shifts of the α-protons were practically unchanged over the temperature range indicates that the amide temperature coefficients are not attributable to a conformational change of the peptide but that they are due to exchange rates associated or not with hydrogen bonding.

Molecular Modeling

Following the molecular modeling procedure described above, 27 of the 35 hCGRP

structures and 29 of the 35 hCGRP 8-37 structures converged onto an identical folding pattern of the N-terminal segment (residues 1-16 and 8-16 for hCGRP and hCGRP 8-37, respectively). The calculated backbone RMSD values in that segment were 1.47 Å and 1.33 Å for the converging structures of hCGRP and hCGRP 8-37, respectively. The lowest energy structures for each peptide are displayed in Figure 4. In all final structures, no NOE distance violations exceeded 0.5 Å. Figure 5 displays ten superimposed structures of the N-terminal segment of both peptides.

The N-terminal hCGRP structure consists of a fairly rigid disulfide-bonded ring followed by two helical segments, Val^{8}-Leu^{12} and Ala^{13} -Leu^{16} , running in opposite directions. The intermediate amide proton temperature coefficients (Figure 2) and the weak secondary shifts (Figure 3) found in that segment are consistent with the fact that this helical segment is not completely regular. Very low amide temperature coefficients were measured in the disulfide-bonded ring (Figure 2) and correspond to hydrogen bonds in the calculated molecular model. All 27 converging structures displayed hydrogen bonding between the amide NH of Thr^{4} and the side chain carbohyl of Asn^{3} as well as between the amide NH of Thr^{6} and the amide carbonyl of Thr^{4} . Twenty-four of the 27 converging structures presented hydrogen bonding between the amide NH of Cys^{7} and the amide carbonyl of Cys^{2} as well as between the amide NH of Val^{8} and both the amide carbonyl of Cys^{2} and the sulfur of Cys^{7} . These two latter hydrogen bonds must exert a strong stabilizing effect on the helix structure.

The C-terminal segment of hCGRP presents little regular structure although several slow exchanging amide protons are present (Figure 2A). The temperature coefficients of these protons are however higher than those of the disulfide-bonded cycle. Among those is the slow exchanging amide proton of Gly^{21} which corresponds to a γ-turn between the amide NH of Gly^{21} and the amide carbonyl of Ser^{19} observed in all converging structures. In addition, hydrogen bonding was observed for 12 structures between the amide NH of Asn^{31} and the carbonyl of Val^{28}, in seven structures between the amide NH of Phe^{27} and the backbone carbonyl of Asn^{26} as well as between the amide NH of Val^{32} and the

backbone carbonyl of Asn^{31}. The hydrogen bond between Val^{28} and Asn^{31} does not give rise to a regular β-turn, as evidenced from the measurement of the torsion angles.

For hCGRP 8-37, removal of the disulfide-bonded ring destabilizes the helix and gives rise to a molecule devoid of regular secondary structure, in agreement with the secondary chemical shifts (Figure 3). However, definite hydrogen bonding is present in two segments of the molecule. The slow exchanging amide proton of Leu^{15} (Figure 4B) corresponds to hydrogen bonding between the amide NH of Leu^{15} and the amide carbonyl of His^{10} in all converging structures. This leads to folding in this part of the molecule (Figures 4 and 5). In addition, hydrogen bonding is observed for 17 structures between the amide NH of Asn^{31} and the amide carbonyl of Val^{28}, similarly to hCGRP. This effect is a result of the presence of a proline residue in position 29.

Our hCGRP structure is similar to the structure reported previously for the same peptide in aqueous: TFE solution at low pH (3) except that the α-helix segment is less regular in our case. Different NH temperature coefficients support that difference. Both models display a γ-turn for Ser^{19}-Gly^{21} but additional hydrogen bonding is present in our model. The stronger helix stabilizing effect of TFE is probably responsible for the structural differences between the two models. The comparison of the three-dimensional structures of hCGRP and hCGRP 8-37 allow to demonstrate the importance of hydrogen bonding between the ring and the helix portion to stabilize the helix. This interaction was not characterized previously.

It appears that the presence of the N-terminal loop might play a major role for peptide activity with $CGRP_1$ receptors and discriminate between agonistic and antagonistic activity. This conclusion is consistent with the results of a structure-activity study which have demonstrated that the α-helical character is essential for activity (18). A circular dichroism study has also concluded that the helical character is reduced in hCGRP 8-37 relative to hCGRP (17). The recent synthesis of an analogue with a C-terminal disulfide-bonded loop displaying agonist activity also supports the importance of the ring for activity (9). On the other hand, our NMR results do not display any connectivity between the N-

and C-terminal portions of the chain. Reports of the importance of both the N- and C-terminal groups for biological activity (18, 19, 21, 25) would suggest that the two terminal ends might be in close proximity, although this has never been demonstrated directly. Development of new analogs and peptidomimetics, and their characterization by NMR, should help in the assessment of the importance of these structural elements for biological activity.

ACKNOWLEDGMENTS

The authors are grateful to Dr. Phan Viet Minh Tan, Université de Montréal for his help with the acquisition of NMR spectra and to Mr. Alain Larocque and Mrs. Anne-Marie Laberge for their help with processing and analysis of the spectra. This work was supported by grants from the Medical Research Council of Canada and the Institut National de la Recherche Scientifique.

REFERENCES

1. **Amara, S.G., J.L. Arriza, S.E. Leff, L.W. Swanson, R.M. Evans and M.G. Rosenfeld.** 1985. Expression in brain of a messenger RNA encoding a novel neuropeptide homologous to calcitonin gene-related peptide. Science *229*: 1094-1097.
2. **Blanco, F.J., J. Herranz, C. Gonzalez, M.A. Jimenez, M. Rico, J. Santoro and J.L. Nieto.** 1992. NMR chemical shifts: a tool to characterize distortions of peptide and protein helices. J. Am. Chem. Soc. *114*: 9676-9677.
3. **Breeze, A.L., T.S. Harvey, R. Bazzo and I.D. Campbell.** 1991. Solution structure of human calcitonin gene-related peptide by ^{1}H NMR and distance geometry with restrained molecular dynamics. Biochemistry *30*: 575-582.
4. **Bundi, A., C. Grathwohl, J. Hochmann, R.M. Keller, G. Wagner and K. Wüthrich.** 1975. Proton NMR of the protected tetrapeptides TFA-Gly-Gly-L-*X*-L-Ala-OCH_3, where X stands for one of the 20 common amino acids. J. Magn. Reson. *18*: 191-198.
5. **Chiba, T., A. Yamaguchi, T. Yamatani, A. Nakamura, T. Morishita, T. Inui, M. Fukase, T. Noda and T. Fujita.** 1989. Calcitonin gene-related peptide receptor antagonist human CGRP-(8-37). Am. J. Physiol. *256*: E331-E335.
6. **Dennis, T., A. Fournier, A. Cadieux, F. Pomerleau, F.B. Jolicoeur, S.A. St-Pierre and R. Quirion.** 1990. hhCGRP 8-37, a calcitonin gene-related peptide antagonist revealing calcitonin gene-related peptide receptor heterogeneity in brain and periphery. J. Pharmacol. Exp. Therap. *254*: 123-128.
7. **Dennis, T., A. Fournier, S.A. St-Pierre and R. Quirion.** 1989. Structure-activity profile of calcitonin gene-related peptide in peripheral and brain tissues. Evidence for receptor multiplicity. J. Pharmacol. Exp. Therap. *251*: 718-725.
8. **Englander, S.W.** 1987. Main-chain-directed strategy for the assignment of ^{1}H NMR spectra of proteins. Biochemistry *26*: 5953-5958.

9. **Hakala, J.M.L., T. Valo, S. Vihavainen, J. Hermonen, P. Heino, M. Halme and A.M. Koskinen.** 1994. Constrained analogues of the calcitonin gene-related peptide. Biochem. Biophys. Res. Commun. *202*: 497-503.

10. **Han, S.P., L. Naes and T.C. Westfall.** 1990. Inhibition of periarterial nerve stimulation-induced vasodilation of the mesenteric arterial bed by CGRP (8-37) and CGRP receptor desensitization. Biochem. Biophys. Res. Commun. *168*: 786-791.

11. **Holzer, P.** 1994. Calcitonin gene-related peptide, p. 493-523. *In* J.H. Walsh and G.J. Dockray (Eds.), Gut peptides, Biochemistry and Physiology. Raven Press, New York.

12. **Hubbard, J.A.M., S.R. Martin, L.C. Chaplin, C. Bose, S.M. Kelly and N.C. Price.** 1991. Solution structures of calcitonin-gene-related peptide analogues of calcitonin-gene-related peptide and amylin. Biochem. J. *275*: 785-788.

13. **Ishida-Yamamoto, A. and M. Tohyama.** 1989. Calcitonin gene-related peptide in the nervous system. Prog. Neurobiol. *33*: 335-386.

14. **Lynch, B. and E.T. Kaiser.** 1988. Biological properties of two models of calcitonin gene related peptide with idealized amphiphilic α-helices of different lengths. Biochemistry *27*: 7600-7607.

15. **Maggi, C.A., T. Chiba and S. Giuliani.** 1991. Human α-calcitonin gene-related peptide-(8-37) as an antagonist of exogenous and endogenous calcitonin gene-related peptide. Eur. J. Pharmacol. *192*: 85-88.

16. **Manning, M.C.** 1989. Conformation of the α form of human calcitonin gene-related peptide (CGRP) in aqueous solution as determined by circular dichroism spectroscopy. Biochem. Biophys. Res. Commun. *160*: 388-392.

17. **Mimeault, M., S.A. St-Pierre and A. Fournier.** 1993. Conformational characterization by circular-dichroism spectroscopy of various fragments and analogs of calcitonin gene-related peptide. Eur. J. Biochem. *213*: 927-934.

18. **Morita, K., I. Kato, T. Uzawa, M. Hori, M. and T. Noda.** 1989. Structure-activity relationship of calcitonin gene related peptide. Hormone Metab. Res. *21*: 666-668.

19. **Noda, T., K. Morita, T. Uzawa, H. Kinoshita and M. Hori.** 1987. Synthesis of an analog of human calcitonin gene related peptide, [Asu]-h-CGRP. Experientia *43*: 890-892.

20. **O'Connell, J.P., S.M. Kelly, D.P. Raleigh, J.A.M. Hubbard, N.C. Price, C.M. Dobson and B.J. Smith.** 1993. On the role of the C-terminus of α-calcitonin-gene-related peptide (αCGRP). The structure of des-phenylalaninamide^{37}-αCGRP and its interaction with the CGRP receptor. Biochem. J. *291*: 205-210.

21. **Otaka, A., N. Fujii, S. Funakoshi, I. Yamamoto, K. Torizuka, T. Noda, K. Morita and H. Yajima.** 1986. Studies on peptides. CXLII. Synthesis of des-1-Ala-des-alpha-amino-human calcitonin gene-related peptide. Chem. Pharm. Bull. *34*: 3915-3918.

22. **Poyner, D.R.** 1992. Calcitonin gene-related peptide: multiple actions, multiple receptors. Pharmac. Ther. *56*: 23-51.

23. **Rosenfeld, M.G., J.J. Mermod, S.G. Amara, L.W. Swanson, P.E. Sawchenko, J. Rivier, W.W. Vale and R.M. Evans.** 1983. Production of a novel neuropeptide encoded by the calcitonin gene via tissue-specific RNA processing. Nature *304*: 129-135.

24. **Steenbergh, P.H., J.W. Höppener, J. Zandberg, C.J.M. Lips and H.S. Jansz.** 1985. A second human calcitonin/CGRP gene. FEBS Lett. *183*: 403-407.

25. **Thiebaud, D., T. Akatsu, T. Yamashita, T. Suda, T. Noda, R.E. Martin, A.E. Fletcher and T.J. Martin.** 1991. Structure-activity relationships in calcitonin gene-related peptide: cyclic AMP response in a preosteoblast cell line (KS-4). J. Bone Miner. Res. *6*: 1137-1142.

26. **Wishart, D.S., B.D. Sykes and F.M. Richards.** 1992. The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. Biochemistry *31*: 1647-1651.

27. **Wüthrich, K.** 1986. *NMR of Proteins and Nucleic Acids*, Wiley, New York.

Table 1. Proton Chemical Shifts (600 MHz) of hCGRP in DMSO-d_6

Residue	Chemical Shift (ppm)				
	NH	αH	βH	γH	Others
Ala-1	8.00	3.83	1.30		
Cys-2	8.83	4.58	2.96,3.30		
Asp-3	8.76	4.62	2.60,2.66		
Thr-4	7.16	4.42	4.13	1.09	
Ala-5	8.63	3.95	1.31		
Thr-6	7.34	4.33	4.02	1.02	
Cys-7	7.66	4.15	3.12,3.26		
Val-8	7.38	4.15	1.94	0.76,0.81	
Thr-9	8.02	4.12	3.92	1.01	
His-10	8.06	4.56	2.94,3.10		
Arg-11	8.05	4.20	1.67	1.53	δCH_2 3.06
					ϵNH 7.54
Leu-12	7.88	4.26	1.45	1.58	δCH_3 0.78,0.94
Ala-13	7.94	4.23	1.19		
Gly-14	8.09	3.69,3.73			
Leu-15	8.05	4.27	1.45	1.58	δCH_3 0.79,0.91
Leu-16	7.96	4.27	1.44	1.57	δCH_3 0.93,0.98
Ser-17	7.93	4.25	3.53,3.71		
Arg-18	7.95	4.30	1.70	1.45	δCH_2 3.04
					ϵNH 7.48
Ser-19	7.79	4.28	3.53,3.74		

Gly-20	7.99	3.62			
Gly-21	8.03	3.75			
Val-22	7.81	4.23	1.92	0.75,0.81	
Val-23	7.98	4.19	1.93	0.84,0.87	
Lys-24	7.87	4.19	1.65	1.20	δCH_2 1.40
					ϵCH_2 2.74
					ζNH_2 7.64
Asn-25	8.04	4.52	2.37,2.51		δNH_2 6.94,7.40
Asn-26	8.09	4.42	2.28,2.42		δNH_2 6.81,7.31
Phe-27	7.90	4.42	2.85,2.99		ArH 7.15
Val-28	7.85	4.11	1.91	0.77,0.80	
Pro-29		4.46	1.85,2.01	1.76	δCH_2 3.51,3.68
Thr-30	7.99	4.18	3.97	1.03	
Asn-31	7.99	4.59	2.50		δNH_2 7.05,7.69
Val-32	7.69	4.09	1.99	0.81,0.91	
Gly-33	8.17	3.71			
Ser-34	7.86	4.23	3.48,3.58		
Lys-35	8.04	4.24	1.61	1.24	δCH_2 1.42
					ϵCH_2 2.72
					ζNH_2 7.64
Ala-36	7.88	4.31	1.09		
Phe-37	7.65	4.36	2.82,3.01		ArH 7.15

Table 2. Proton Chemical Shifts (600 MHz) of hCGRP 8-37 in DMSO-d_6

Residue	Chemical Shift (ppm)					
	NH	αH	βH	γH	Others	
Val-8	8.04	3.75	2.03	0.84,0.88		
Thr-9	8.37	4.28	3.93	1.04		
His-10	8.24	4.65	2.93,3.10			
Arg-11	8.14	4.32	1.67	1.46	δCH_2	3.08
					ϵNH	7.59
Leu-12	7.91	4.08	1.56	1.41	δCH_3	0.82,0.87
Ala-13	7.98	4.24	1.21			
Gly-14	8.02	3.68				
Leu-15	8.08	4.24	1.58	1.45	δCH_3	0.82,0.86
Leu-16	7.99	4.30	1.56	1.46	δCH_3	0.82,0.86
Ser-17	7.89	4.31	3.54,3.60			
Arg-18	7.99	4.23	1.72	1.49	δCH_2	3.07
					ϵNH	7.51
Ser-19	7.95	4.29	3.55,3.60			
Gly-20	8.03	3.64				
Gly-21	8.11	3.75				

Val-22	8.12	4.19	1.97	0.84,0.88		
Val-23	7.82	4.12	1.94	0.84		
Lys-24	7.88	4.19	1.68	1.29	δCH_2	1.50
					ϵCH_2	2.74
					ζNH_2	7.63
Asn-25	8.06	4.52	2.39,2.52		δNH_2	6.96,7.41
Asn-26	8.10	4.45	2.32,2.46		δNH_2	6.83,7.33
Phe-27	7.91	4.45	2.85,3.02		ArH	7.17,7.23
Val-28	7.87	4.23	1.96	0.87,0.89		
Pro-29		4.48	1.79,1.90	2.03	δCH_2	3.52,3.70
Thr-30	7.88	4.19	3.98	1.03		
Asn-31	8.00	4.60	2.48,2.54		δNH_2	6.98,7.43
Val-32	7.70	4.10	2.01	0.87,0.89		
Gly-33	8.18	3.74				
Ser-34	7.82	4.28	3.54,3.58			
Lys-35	8.07	4.25	1.64	1.29	δCH_2	1.50
					ϵCH_2	2.74
					ζNH_2	7.63
Ala-36	7.88	4.14	1.12			
Phe-37	7.68	4.35	2.82,3.00		ArH	7.07,7.23

FIGURES

Figure 1. Diagonal plot of interresidual connectivities observed for (a) hCGRP and (b) hCGRP 8-37 in DMSO-d_6. A filled square indicates that at least one backbone-backbone connectivity was observed between the two residues, a shaded square indicates that at least one backbone-sidechain connectivity was observed and an open square indicates that only sidechain-sidechain connectivities were observed.

Figure 2. Summary of sequential and medium range NOESY connectivities indicative of secondary structure which were observed for (a) hCGRP and (b) hCGRP 8-37. Filled and open squares correspond to low and intermediate temperature coefficients of the amide protons, respectively.

Figure 3. Secondary chemical shifts (SCS) for the α-protons of hCGRP and hCGRP 8-37 in DMSO-d_6 as a function of residue number. Negative shifts of -0.1 and more are indicative of helical structure.

Figure 4. Lowest-energy structures of (a) hCGRP and (b) hCGRP 8-37 following molecular modeling by simulated annealing and energy minimization calculations using NOE constraints from NMR.

Figure 5. Superimposition of the N-terminal residues for the ten lowest energy structures of (a) hCGRP and (b) hCGRP 8-37 following molecular modeling calculations using NOE constraints from NMR. Backbone-backbone hydrogen bonding (2.2-2.5 Å) corresponding to amide protons with low temperature coefficients is illustrated by dotted lines.

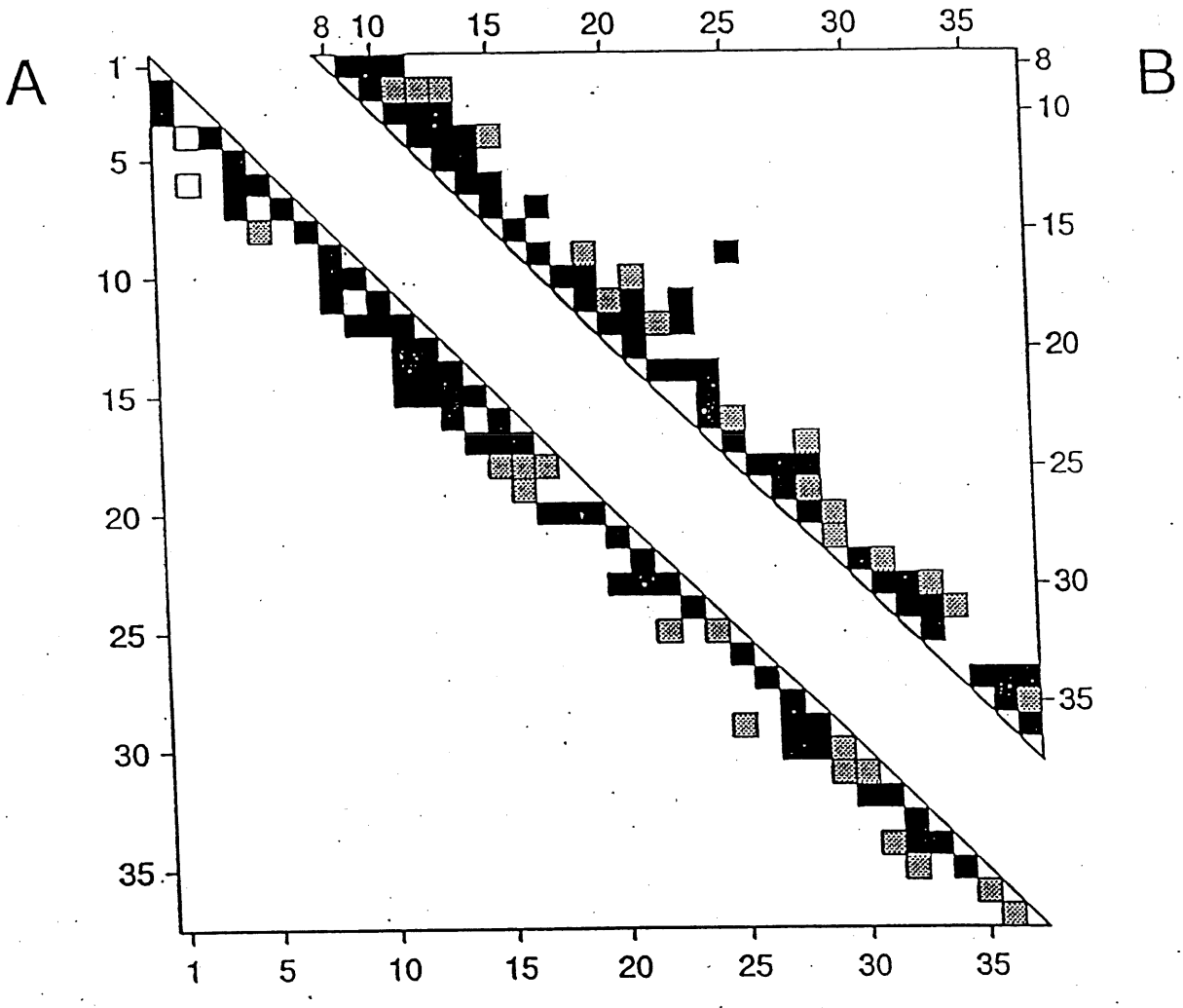


Fig. 1

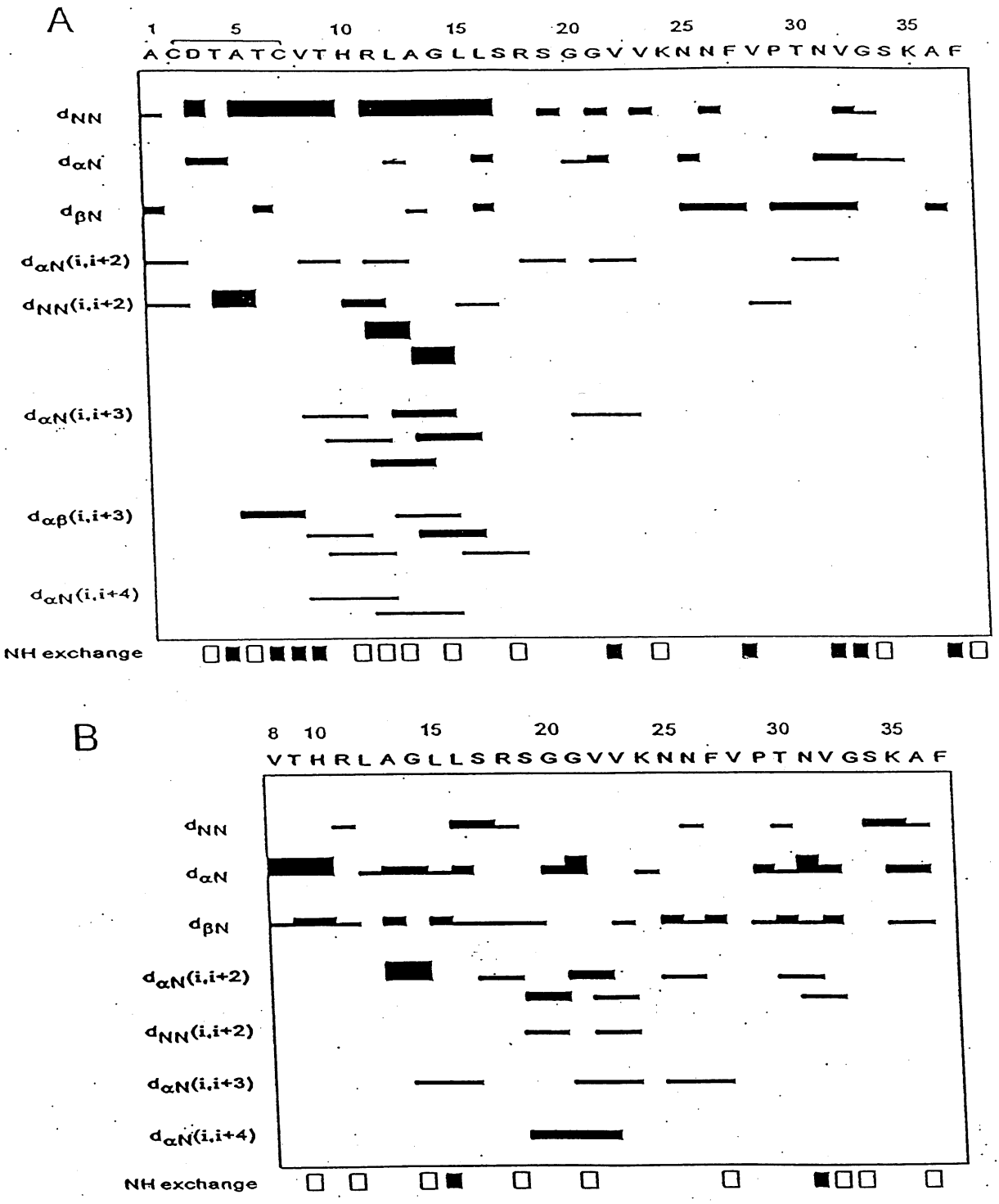

A
1 5 10 15 20 25 30 35
ACDTATCVTHRLAGLLSRSGGVVKNNFVPTNVGSKAF
dNN
dαN
dβN
dαN(i,i+2)
dNN(i,i+2)
dαN(i,i+3)
dαβ(i,i+3)
dαN(i,i+4)
NH exchange
B
8 10 15 20 25 30 35
VTHRLAGLLSRSGGVVKNNFVPTNVGSKAF
dNN
dαN
dβN
dαN(i,i+2)
dNN(i,i+2)
dαN(i,i+3)
dαN(i,i+4)
NH exchange

Fig. 2

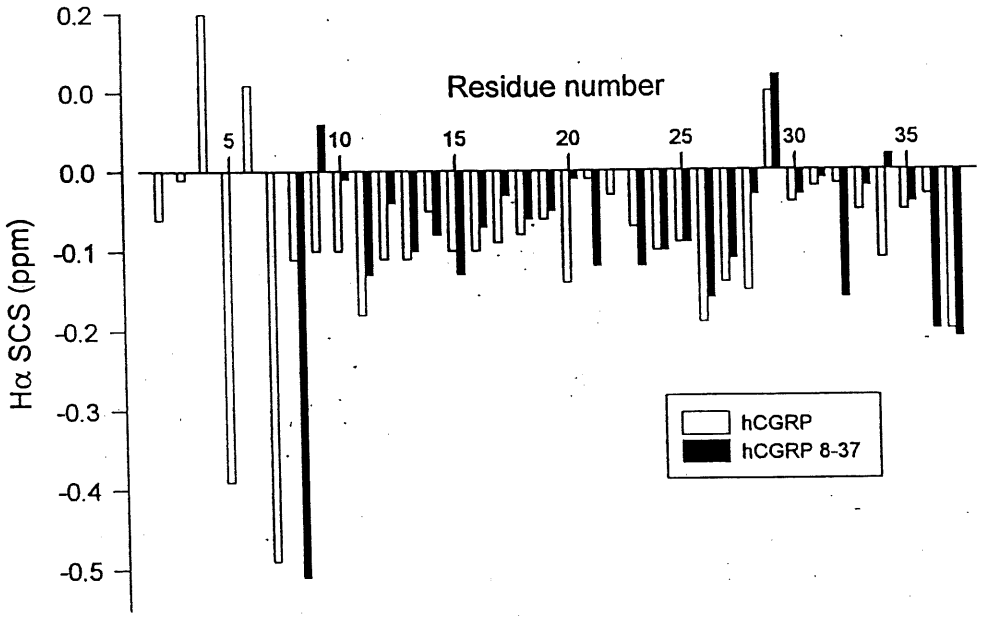
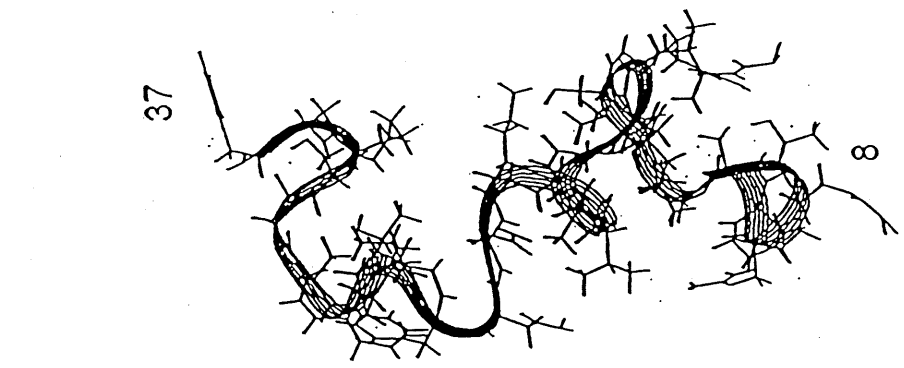
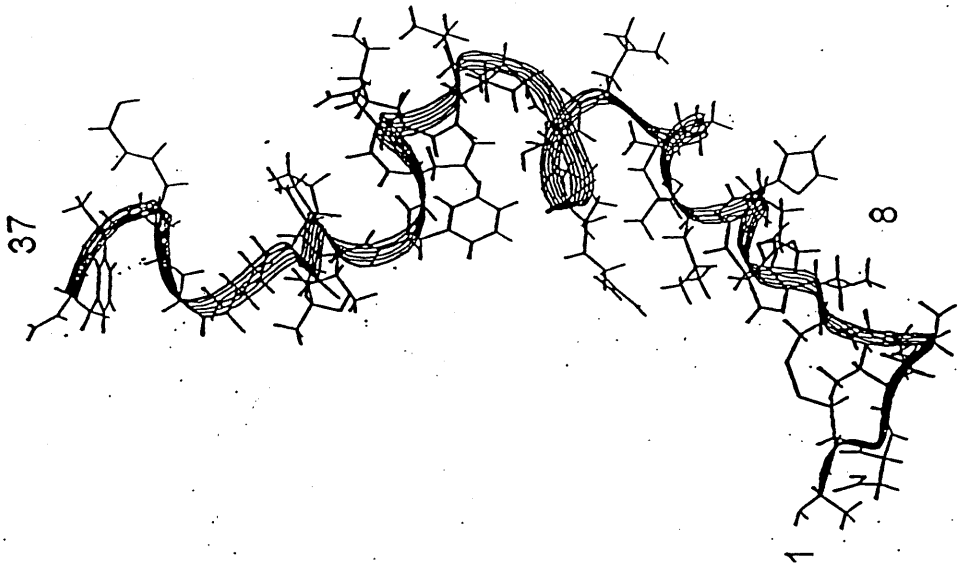

0.2
0.0
0.0
-0.1
-0.2
-0.3
-0.4
-0.5
Hα SCS (ppm)
Residue number
5
10
15
20
25
30
35
hCGRP
hCGRP 8-37

Fig. 3



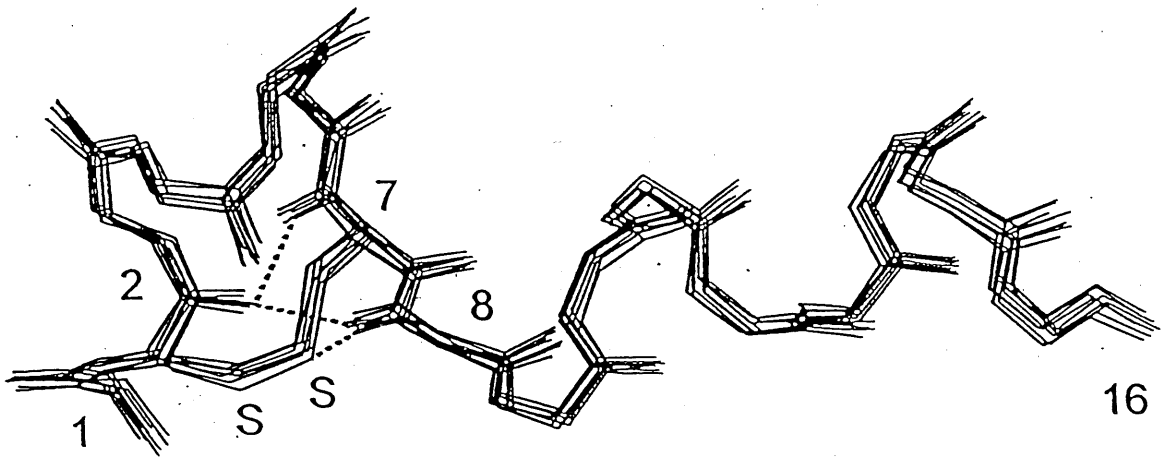
B



A

Fig. 4

A



B

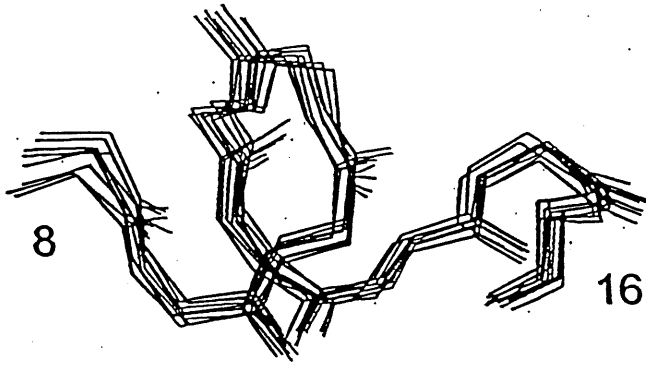


Fig. 5

CHAPITRE III

Structural Effects of the Selective Reduction of Amide Carbonyl Groups in Motilin 1-12 as Determined by NMR†

YVAN BOULANGER,[1] ABDESSLEM KHIAT,[1] YANMIN CHEN,[1]
DIANE GAGNON,[1,2] PIERRE POITRAS,[2] and SERGE ST-PIERRE[1]

[1]*INRS-Santé, Université du Québec, Pointe-Claire, Quebec, Canada, and* [2]*Centre de Recherche Clinique André-Viallet, Hôpital St-Luc, Montreal, Quebec, Canada*

Short title: Structure of Motilin 1-12 Analogs

International Journal of Peptide and Protein Research, vol. 46, 527 (1995)

INTRODUCTION

La motiline est un peptide de 22 résidus qui stimule la motilité de l'estomac et de l'intestin. Le fragment 1-12 montre un effet biologique similaire à celui du peptide natif. La réduction sélective du groupement carbonyle amide pour former les analogues CH_2NH produit une réduction significative de l'activité pour les deux premières positions N-terminales et une perte complète de l'activité pour toutes les autres positions. Le but du travail présenté est d'essayer d'expliquer la réduction ou la perte totale d'activité des analogues CH_2NH en analysant leurs structures tridimensionnelles. Dans un premier temps, l'effet de la température sur le déplacement chimique des protons NH dans dix analogues motiline 1-12, $(CH_2NH)^{1-2}$, $(CH_2NH)^{2-3}$ ... $(CH_2NH)^{10-11}$-motiline 1-12 est étudié afin de localiser les liens hydrogènes. Dans un deuxième temps les structures secondaires et tertiaires sont déterminées pour trois analogues, l'un actif, la motiline 1-12, le deuxième moins actif, la $(CH_2NH)^{1-2}$-motiline 1-12 et le troisième complètement inactif, la $(CH_2NH)^{4-5}$-motiline 1-12. Les résultats montrent que dans tous les analogues, la structure de la région N-terminale (résidus 1-5) est différente de celle de la motiline 1-12, cette dernière étant caractérisée par un lien hydrogène entre le Phe^1 et le Ile^4. La structure de la région C-terminale des analogues est similaire à celle de la motiline 1-12 pour seulement les deux premières positions réduites (1-2 et 2-3), indiquant l'importance du segment C-terminal de la motiline 1-12 et la conservation de la rigidité de la liaison amide pour l'activité biologique.

Pour la réalisation de ce travail, j'ai acquis toutes les bases théoriques relatives à la modélisation moléculaire au département de chimie de l'Université de Sherbrooke, j'ai effectué le travail de recherche et de rédaction de la partie modélisation moléculaire et une partie de la RMN à l'INRS-Santé. Les mesures RMN ont été faites à l'université de Montréal.

ABSTRACT

Motilin is a 22-residue peptide stimulating stomach and intestinal motility. The motilin 1-12 fragment displays biological effects similar to the native peptide. Selective reduction of the amide carbonyl groups to form $\Psi[CH_2NH]$ analogs leads to a significant reduction in activity for the first two N-terminal positions and to a complete loss of activity for all other positions. The structures of motilin 1-12 and ten reduced analogs were investigated using the temperature dependence of the amide NH chemical shifts. In all the analogs, the structure of the N-terminal region (residues 1-5) was different from the structure of motilin 1-12, which is characterized by hydrogen bonding between Phe^1 and Ile^4. The structure of the C-terminal region of analogs was similar to the structure of motilin 1-12 for the first two reduction positions only (1-2 and 2-3), indicating that the C-terminal portion of motilin 1-12 is more critical for biological activity. Complete structural characterizations of motilin 1-12, $\Psi[CH_2NH]^{1-2}$- and $\Psi[CH_2NH]^{4-5}$-motilin 1-12 were performed by two-dimensional NMR spectroscopy and molecular modeling. The structural features observed confirm the differences based on the temperature dependence of the amide NH chemical shifts. These results demonstrate that conservation of the amide bond rigidity is essential for the activity of non-hydrolyzable analogs.

Key Words: enzymatic cleavage protection analogs; molecular modeling; peptide conformation; structure-activity relationship; two-dimensional NMR.Abbreviations: DMSO, dimethylsulfoxide; DOPC, 1,2-dioleoyl-*sn*-glycero-3-phosphocholine; DOPG, 1,2-dioleoyl-*sn*-glycero-3-phosphoglycerol; DQFCOSY, double quantum filtered correlation spectroscopy; FAB, fast atom bombardment; HFP, 1,1,1,3,3,3-hexafluoro-2-propanol; HPLC, high performance liquid chromatography; NOE, nuclear Overhauser enhancement; NOESY, nuclear Overhauser enhancement spectroscopy; RMS, root-mean-square; RMSD, root-mean-square deviation; SDS, sodium dodecyl sulphate; TOCSY, total correlation spectroscopy.

Motilin is a 22-residue peptide hormone which stimulates the contractile activity of the stomach and upper small intestine (1-3). The peptide is synthesized by endocrine cells of the duodenum and jejenum mucosa, it circulates in blood and acts mostly to stimulate the phase III of the migrating motor complex in the interdigestive period (4-6). The motilin receptor was recently solubilized and characterized from rabbit antral smooth muscle tissue (7). Erythromycin A and erythromycin derivatives were found to be motilin receptor agonists (8-10).

The structures of porcine and rabbit motilin have been investigated in 1,1,1,3,3,3-hexafluoro-2-propanol (HFP): water mixtures as well as in 1,2-dioleoyl-*sn*-glycero-3-phosphocholine (DOPC) vesicles, 1,2-dioleoyl-*sn*-glycero-3-phosphoglycerol (DOPG) vesicles and sodium dodecyl sulphate (SDS) micelles using circular dichroism and two-dimensional ^{1}H NMR spectroscopy (11-14). α-Helical structure was found to exist in aqueous solvent and is stabilized by the presence of HFP (11), negatively charged vesicles (DOPG) or micelles (SDS) but not by the presence of neutral DOPC lipid vesicles (13, 14). The solvent dependence of α-helix formation was similar for the motilin 1-16 fragment and the native motilin (14). Analysis of the NMR results in 30% HFP allowed to determine that the α-helix extends from Glu^9 to Lys^{20} and that a wide turn is formed between Pro^3 and Thr^6 (11, 12).

Structure-activity studies on synthetic fragments of motilin demonstrated that the N-terminal segment is particularly important for activity (15-20). C-terminal fragments failed to stimulate any contractile activity (15-17) and an important loss of activity was measured when Phe^1 was substituted or deleted (18, 21). In rabbit duodenal muscle preparations *in vitro*, the N-terminal fragments 1-12 and 1-15 induced a contractile activity comparable to that induced by the native peptide, fragments 1-10 and 1-11 induced significant contractile activity but shorter fragments were inactive (18, 20). However, in the conscious dog, none of these N-terminal fragments influenced the intestinal myoelectrical activity (19).

Selective reduction of the amide groups of motilin 1-12 to form $\Psi[CH_2NH]$ isosteres resistant to enzymatic cleavage resulted in a decrease in contractile response for the analogs reduced in positions 1-2 and 2-3 and in a total loss of activity for all other positions (20). Table 1 presents the ED_{50} values measured for motilin 1-12 and its $\Psi[CH_2NH]$ analogs on the contractile activity of the rabbit duodenal muscle *in vitro* (20). These data show that a reduced activity is retained when the $\Psi[CH_2NH]$ group is in position 1-2 and 2-3 and that all the other analogs are inactive. The contractile response for the $\Psi[CH_2NH]^{1-2}$ and $\Psi[CH_2N]^{2-3}$ analogs is approximately five and ten times smaller, respectively, than for motilin 1-12 (Table 1).

In this study, the three-dimensional structures of these $\Psi[CH_2NH]^{n,n+1}$motilin 1-12 analogs were investigated by NMR spectroscopy and molecular modeling. NMR measurements were performed in DMSO-d_6, a moderately polar solvent stabilizing peptide structures and considered a good mimic of receptor site environments. Our data demonstrate a definite correlation between changes in contractile response *in vitro* and changes in C-terminal three-dimensional structure occurring around reduction position 3-4.

EXPERIMENTAL PROCEDURES

Peptide synthesis

Porcine motilin 1-12 was synthesized by solid phase peptide synthesis as described previously (20). Briefly, the synthesis was performed with a homemade manual multireactor synthesizer using *tert*-butyloxycarbonyl (Boc) amino acid or amino acid aldehyde derivatives, *p*-methylbenzhydrylamine resin as solid support and benzotriazol-1-yloxy-tris(dimethylamino)phosphonium hexafluorophosphate (BOP) as coupling agent. The peptide was deprotected and cleaved from the resin using hydrofluoric acid (HF) in the presence of m-cresol. The crude peptide was finally purified by preparative C_{18} reverse-phase HPLC.

$\Psi[CH_2NH]^{n,n+1}$motilins 1-12 were synthesized by the same method using Boc-amino acid aldehydes at the appropriate positions (20). The coupling of amino acid aldehydes was performed in the presence of the sodium cyanoborohydride ($NaBH_3CN$) reducing agent. No coupling of the secondary $\Psi[CH_2NH]$ amine occurs during this reaction (22). All peptides were obtained with a purity of more than 98% as monitored by analytical HPLC, amino acid analysis and FAB mass spectrometry.

NMR measurements

NMR samples of motilin 1-12 and $\Psi[CH_2NH]^{n,n+1}$motilin 1-12 analogs (trifluoroacetate salts) were prepared in 0.5 ml of DMSO-d_6 at a concentration of 4.1 mM and transferred to 5-mm NMR tubes. Samples were deoxygenated by bubbling argon gas and then sealed. All NMR experiments were performed on a Bruker ARX 400 spectrometer operating at 400.14 MHz or on a Bruker AMX 600 spectrometer operating at 600.13 MHz (Laboratoire Régional de RMN à Haut Champ, Université de Montréal). One-dimensional ^1H NMR spectra were recorded using 8196 data points at temperatures of 310, 315, 320, 325, 330, 335 and 340 K in order to determine the temperature dependence of the amide chemical shifts. Phase-sensitive double-quantum-filtered correlation spectroscopy (DQFCOSY) spectra (23), total correlation spectroscopy (TOCSY) spectra (24, 25) and nuclear Overhauser enhancement spectroscopy (NOESY) spectra in the time-proportional-phase-incrementation mode with mixing times of 200-400 ms (26) were collected according to standard procedures. All two-dimensional NMR spectra were recorded using 1024 data points and 256 t_1 increments. The NMR spectra were processed using the FELIX 2.30 software (Biosym Technologies, Inc., San Diego CA) operating on a Silicon Graphics Indigo R4000 XS workstation. One-dimensional spectra were zero-filled to 16392 data points and Fourier transformed. Two-dimensional spectra were zero-filled to 2048 data points in each dimension and multiplied by a 30°-shifted sinebell squared window function prior to Fourier transformation.

Sequential resonance assignments were performed using conventional methods (27). Distance constraints were derived from the intensities of the NOESY cross-peaks. NOE intensities were classified as strong, medium and weak corresponding to upperbound constraints of 0.3, 0.4 and 0.5 nm, respectively. When equivalent proton nuclei were present on a same group, a 0.05 nm distance correction was added to the constraint.

Molecular modeling

Molecular modeling calculations were carried out on the Silicon Graphics workstation using software purchased from Biosym Technologies, Inc. (INSIGHT II, DISCOVER and NMRCHITECT version 2.30). In order to sample adequately the conformational space, the extended structure of each peptide, built using the consistent valence forcefield, was first submitted to 500 iterations of conjugate gradient energy minimization with a RMS gradient of 0.001 Å, followed by molecular dynamics simulation using 5 fs steps to generate 10^5 conformations sampled at every 100 fs. Cluster analysis of these conformations on the basis of interatomic distances allowed to classify them into seven families of conformations. The lowest energy conformation in each family was then selected. Each conformation was energy minimized for 1000 iterations using the steepest descent algorithm. The NMR constraints were then introduced and the molecules were heated from 300 to 1000K over the course of 10 ps, then cooled over a period of 4 ps to 500K and finally cooled to 300K over a period of 4 ps. Following dynamics, the peptides were subjected to 2000 additional iterations of conjugate gradients energy minimization with a RMS gradient of 0.001 Å. Five structures were generated for each of the seven starting conformations.

RESULTS

Exchange rates of amide protons

The rates of exchange of the amide protons were estimated from the temperature dependence of their ^{1}H NMR chemical shifts (Table 2). Temperature coefficients higher than -4 x 10^{-3} ppm/K in DMSO are indicative of restricted mobility for the hydrogen atom, which usually results from hydrogen bonding. The values of Table 2 indicate that most amide hydrogens are not involved in hydrogen bonding. The absence of values at certain positions is due to the absence of observable signal, this occurring mostly in the vicinity of the $\Psi[CH_2NH]$ replacement.

Comparison of the temperature coefficients of the $\Psi[CH_2NH]$ analogs with those of the native peptide provides an indication of the degree of structural change resulting from the reduction of the amide bond. When comparing the temperature coefficients of the $\Psi[CH_2NH]^{1-2}$ and $\Psi[CH_2N]^{2-3}$ analogs to those of the native peptide, nearly identical values are observed for the amide protons of residues 6-12 (Table 2). This suggests that the structure of the C-terminal part of these molecules is identical, which is not the case for any of the other analogs. On the other hand, the temperature coefficients of practically all analogs are different from those of the native peptide in the N-terminal portion (residues 1-5).

Two-dimensional NMR results

Two-dimensional DQFCOSY, TOCSY and NOESY spectra were recorded for motilin 1-12, $\Psi[CH_2NH]^{1-2}$ motilin 1-12 and $\Psi[CH_2NH]^{4-5}$ motilin 1-12 in DMSO-d_6. Complete resonance assignments were realized for each peptide using DQFCOSY and TOCSY spectra for the assignment of the spin systems of specific residues and NOESY spectra for sequential assignments (Fig. 1). A comparison of the ^{1}H chemical shifts for the three molecules is given in Table 3. The chemical shifts are very similar for all three peptides except in the vicinity of the reduced amide bond, suggesting structural differences in these regions. Signals were not observed for the C-terminal amide

groups of motilin 1-12 and $\Psi[CH_2NH]^{1-2}$motilin 1-12 as well as in the vicinity of the $\Psi[CH_2NH]$ group for $\Psi[CH_2NH]^{4-5}$motilin 1-12.

The NOE distance constraints for motilin 1-12 comprised 42 intra-residue and 54 inter-residue connectivities which include 45 sequential and 9 medium range (7 i,i+2, 1 i,i+3 and 1 i,i+4) connectivities. In the case of $\Psi[CH_2NH]^{1-2}$motilin 1-12, 55 intra-residue and 34 inter-residue connectivities, including 29 sequential and 5 i,i+2 connectivities, were observed. For $\Psi[CH_2NH]^{4-5}$motilin 1-12, 38 intra-residue and 27 inter-residue connectivities (all sequential) were observed. In no case were any long range connectivities (more than four residues apart) detected.

Backbone dihedral angle constraints were derived from the $^3J_{NH\alpha H}$ coupling constant values (Table 4). The coupling constants are nearly identical for motilin 1-12 and $\Psi[CH_2NH]^{1-2}$motilin 1-12 but differ for $\Psi[CH_2NH]^{4-5}$motilin 1-12. Eleven, ten and seven dihedral angle constraints were obtained for motilin 1-12, $\Psi[CH_2NH]^{1-2}$motilin 1-12 and $\Psi[CH_2NH]^{4-5}$motilin 1-12, respectively. All values are inconsistent with any type of regular secondary structure formation but could be due to conformational averaging in a flexible structure (28).

Molecular modeling

Following the energy minimization and molecular dynamics procedures described above, 31, 30 and 27 of the 35 final structures of motilin 1-12, $\Psi[CH_2NH]^{1-2}$motilin 1-12 and $\Psi[CH_2NH]^{4-5}$motilin 1-12, respectively, converged onto an identical folding pattern. None of these structures displayed NOE distance violations of more than 0.05 nm and the calculated backbone RMSD values were 0.13±0.03 nm, 0.14±0.03 nm and 0.17±0.05 nm for motilin 1-12, $\Psi[CH_2NH]^{1-2}$motilin 1-12 and $\Psi[CH_2NH]^{4-5}$motilin 1-12, respectively.

The converging structures of the three molecules were analyzed in terms of regular secondary structure. In the motilin 1-12 molecule, hydrogen bonding is observed in 17 out of 35 structures between the amide hydrogen of Phe^1 and the carbonyl oxygen of Ile^4 (Fig. 2a), leading to the formation of a 14-membered

pseudocycle. This hydrogen bond is suggestive of a β-turn but the measured torsion angles (Φ,Ψ) for Val[2] and Pro[3] do not correspond to any regular type of β-turn, the existence of a hydrogen bond between residue i and residue i+3 being insufficient to identify a β-turn (29, 30). In the $\Psi[CH_2NH]^{1-2}$motilin 1-12 molecule, no hydrogen bonding between the Phe[1] and Ile[4] residues exists. This is in agreement with the larger negative temperature coefficient calculated for $\Psi[CH_2NH]^{1-2}$motilin 1-12 than for motilin 1-12 (Table 2). In the case of $\Psi[CH_2NH]^{4-5}$motilin 1-12, Phe[1] -Ile[4] folding is not observed but hydrogen bonding between the amide hydrogen of Phe[1] and the carbonyl oxygen of Val[2] is found in 15 of the 35 structures (Fig. 2b). Moreover, hydrogen bonding was found in 25 of the 35 structures between the C-terminal amide hydrogens and the carbonyl oxygen of Gln-11 (Fig. 2c). These hydrogen bonding patterns are consistent with the low temperature coefficients observed for the Phe[1] and C-terminal amide hydrogens of $\Psi[CH_2NH]^{4-5}$motilin 1-12 (Table 2).

The general structures of both motilin 1-12 and $\Psi[CH_2NH]^{1-2}$motilin 1-12 are similar and present an analogous folding pattern in the C-terminal part but differ in the N-terminal part (Fig. 3). However, the structure of $\Psi[CH_2NH]^{4-5}$motilin 1-12 does not present any short range folding and therefore differs from the two other molecules. For all three molecules, the Gly[8]-Gln[11] region is the best defined.

DISCUSSION

These conformational results based on different NMR parameters (NH exchange rates, chemical shifts, $^3J_{NH\alpha H}$ coupling constants, NOE distances) measured for motilin 1-12 and its analogs demonstrate that reduction of a single backbone carbonyl group leads to significant structural effects. Reduction of the carbonyl group in position 1-2 causes structural changes in the N-terminal segment mainly characterized by the disappearance of the hydrogen bond formation between Phe[1] and Ile[4]. Reduction of the carbonyl group in position 4-5 leads to more significant structural modifications in both

the N- and C-terminal parts of the molecule. Hydrogen bonding is now forming between Phe^1 and Val^2 as well as between the C-terminal amide and Gln^{11} (Fig. 2c). The general structure of $\Psi[CH_2NH]^{4-5}$motilin 1-12 does not display wide turns, as is the case of motilin 1-12 and $\Psi[CH_2NH]^{1-2}$motilin 1-12 (Fig. 3).

As expected for peptides of this size in solution, their structure is relatively flexible in solution. This is indicated by the large temperature coefficients for most NH protons (Table 1), by the average $^3J_{NH\alpha H}$ values (Table 4), by the low number of NOE constraints and by the relatively large RMSD values following molecular modeling. However, the peptide dynamics does not prevent the identification of clear structural features as described above. The structural variability is illustrated in Fig. 3 for 10 low energy conformers of each molecule. Previous NMR studies reported for motilin 1-22 also reported significant mobility in the N-terminal region (11, 12). Our determination of coupling constants and temperature coefficients, in addition to NOE distance data, allows us to obtain a better defined structure than previously reported for the same portion in the whole peptide.

These NMR-derived structural results are self-consistent and can be easily correlated with biological activity (20). Only when carbonyl groups were reduced in position 1-2 or 2-3 was there significant biological activity (Table 1). Examination of the temperature coefficients in Table 2 indicates that the C-terminal (Thr^6-Arg^{12}) temperature coefficients of $\Psi[CH_2NH]^{1-2}$- and $\Psi[CH_2N]^{2-3}$motilin 1-12 are similar to those of the native motilin 1-12. All other analogs display different temperature coefficients in the C-terminal region and are therefore structurally different from the native motilin 1-12, as is illustrated by the detailed structure determined for $\Psi[CH_2NH]^{4-5}$motilin 1-12 (Fig. 3). On the other hand, the temperature coefficients of all reduced analogs in the N-terminal part (Phe^1-Phe^5) are different from those of motilin 1-12 and biological activity is not always destroyed. Therefore, the structure of this region appears to be less critical for activity, although the activity is reduced by 5-10 fold upon reduction of the carbonyl group in positions 1-2 and 2-3 (Table 1). In spite of the formation of hydrogen bonding involving Phe^1, the N-terminal segment could be fairly

mobile and increased mobility resulting from amide bond reduction might not affect strongly the biological activity. However, the important loss of activity when the Phe^1 residue is substituted or removed suggests that the nature of the side chain is critical in that segment (18, 21). On the other hand, the rigidity of the C-terminal segment appears to be an essential element for activity since the reduced amide bonds destroy all activity. The structural importance of the C-terminal region of motilin 1-12 is in agreement with the fact that shorter N-terminal motilin fragments were found to produce strongly reduced or no contractile activity (18, 20).

To our knowledge, this is the first report of the structural effects caused by the reduction of amide bonds in peptides of this size. Replacement of the amide carbonyl group by a methylene group destroys the rigidity associated with the planar amide bond and adds a positive charge on the $\Psi[CH_2NH]$ moiety. This leads to more flexibility and modified interactions in the vicinity of the aminomethylene group, affecting more or less the overall molecular structure depending on the position of the reduced carbonyl group. In the case of motilin 1-12, the first two N-terminal reductions produced localized effects but all the others affected the overall molecular structure. Therefore, the desirable effects of protection against enzymatic cleavage provided by the methylene group produce structural disturbances resulting from the loss of planarity (and rigidity) of the amide bond, affecting or destroying biological activity. Based on these results, conservation of the amide bond rigidity seems to be an essential element for the design of biologically active non-hydrolyzable peptide analogs.

REFERENCES

1. Brown, J.C., Cook, M.A., & Dryburgh, J.R. (1972) *Can. J. Biochem.* **51**, 533-537.
2. Itoh, Z. (1990) *Motilin*, Academic Press, San Diego.
3. Poitras, P. (1994) in *Gut Peptides: Biochemistry and Physiology* (Walsh, J.H., & Dockray, G.J., eds.), pp. 261-304, Raven Press, New York.
4. Poitras, P. (1984) *Gastroenterology* **87**, 913-928.
5. Usellini, L., Buchan, A.M.J., Polak, J.M., Capella, C., Cornaggia, M., & Solcia, E. (1984) *Histochemistry* **81**, 363-368.
6. Vantrappen, G., & Peeters, T.L. (1989) in *Handbook of Physiology*, vol. II, Section G, pp. 545-558, American Physiology Society, Bethesda.
7. Depoortere, I., Peeters, T.L., & Vantrappen, G. (1993) *J. Receptor Res.* **13**, 903-923.
8. Peeters, T., Matthijs, G., Depoortere, I., Cachet, T., Hoogmartens, J., & Vantrappen, G. (1989) *Am. J. Physiol.* **257**, G470-G474.
9. Depoortere, I., Peeters, T.L., & Vantrappen, G. (1990) *Peptides* **11**, 515-519.
10. Koga, H., Sato, T., Tsuzuki, K., Onoda, H., Kuboniwa, H., & Takanashi, H. (1994) *Bioorg. Med. Chem. Lett.* **4**, 1347-1352.
11. Khan, N., Gräslund, A., Ehrenberg, A., & Shriver, J. (1990) *Biochemistry* **29**, 5743-5751.
12. Edmondson, S., Khan, N., Shriver, J., Zdunek, J., & Gräslund, A. (1991) *Biochemistry* **30**, 11271-11279.
13. Backlund, B.-M., & Gräslund, A. (1992) *Biophys. Chem.* **45**, 17-25.
14. Backlund, B.-M., Wikander, G., Peeters, T.L., & Gräslund, A. (1994) *Biochim. Biophys. Acta* **1190**, 337-344.
15. Segawa, T., Nakano, M., Kai, Y., Kawatani, H., & Yajima, H. (1976) *J. Pharm. Pharmacol.* **28**, 650-651.

16. Yajima, H., Kai, Y., Ogawa, H., Kubota, M., Mori, Y., & Kojami, K. (1977) *Gastroenterology* **72**, 793-796.
17. Peeters, T.L., Bormans, V., Depoortere, I., Matthijs, G., Kitagawa, K., & Vantrappen, G. (1988) *Biomed. Res.* **9**, 361-366.
18. Poitras, P., Gagnon, D., & St-Pierre, S. (1992) *Biochem. Biophys. Res. Commun.* **183**, 36-40.
19. Raymond, M.C., Boivin, M., St-Pierre, S., Gagnon, D., & Poitras, P. (1994) *Regul. Pept.* **50**, 121-126.
20. Miller, P., Gagnon, D., Dickner, M., Aubin, P., St-Pierre, S., & Poitras, P. (1995) *Peptides* **16**, 11-18.
21. Kitagawa, K., Yonoto, K., Kiyama, S., Ando, K., Kawamoto, T., Akita, T., Inoue, A., & Segawa, T. (1985) *Chem. Pharm. Bull.* **33**, 3307-3316.
22. Sasaki, Y., & Coy, D.H. (1987) *Peptides* **8**, 119-121.
23. Derome, A.E., & Williamson, M.P. (1990) *J. Magn. Reson.* **88**, 177-185.
24. Bax, A., & Davis, D.G. (1985) *J. Magn. Reson.* **65**, 355-360.
25. Rance, M. (1987) *J. Magn. Reson.* **74**, 557-564.
26. Bodenhausen, G., Kogler, H., & Ernst, R.R. (1984) *J. Magn. Reson.* **58**, 370-388.
27. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, J. Wiley and Sons, New York.
28. Hoch, J.C., Dobson, C.M., & Karplus, M. (1985) *Biochemistry* **24**, 3831-3841.
29. Zimmerman, S.S., & Scheraga, H.A. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 4126-4129.
30. Chou, P.Y., & Fasman, G.D. (1977) *J. Mol. Biol.* **15**, 135-148.

ACKNOWLEDGMENTS

The authors would like to thank Mrs. Anne-Marie Laberge and Mr. Alain Larocque for the transformation and analysis of NMR spectra, Mr. Louis Senécal for the support of the computer instrumentation as well as Dr. Yongxue Tu and Dr. Phan Viet Minh Tan for their help with the acquisition of NMR spectra. This study was supported by the Institut National de la Recherche Scientifique and by the Medical Research Council of Canada.

TABLE 1

Effect of Motilin 1-12 and Ψ[CH_2NH] Analogs on Contractile Activity of Rabbit Duodenal Wall Muscle[a]

Ψ[CH_2NH] Position on Motilin 1-12	Muscle Contraction EC_{50} ± SEM (nM)
Native	4 ± 2
1-2	39 ± 11
2-3	21 ± 10
3-4	> 1000
4-5	> 10,000
5-6	> 10,000
6-7	> 10,000
7-8	> 10,000
8-9	> 10,000
9-10	> 10,000
10-11	> 1000

[a] From Miller et al., 1995.

TABLE 2

Variation of the ^1H NMR Chemical Shifts vs Temperature for the Amide Protons of Motilin 1-12 and Analogs

		Ψ[CH_2NH] Position									
	NATIVE	1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	9-10	10-11
F1 NH_2	-1.71	-5.07	-5.99	-8.47	-2.54	-2.31	-1.74	-2.94	-4.94	-2.71	-3.07
V2 NH	-6.96		-7.00	-7.09	-7.11	-7.43	-6.89	-7.44	-6.89	-7.39	-7.01
I4 NH	-5.71	-7.18				-5.01	-5.44	-5.52	-6.04	-6.24	-5.84
F5 NH	-5.29	-6.13					-6.04				-5.20
T6 NH	-6.99	-7.14	-8.01	-4.79				-5.14	-5.66	-7.56	-7.11
Y7 NH	-4.94	-4.99	-4.60	-4.39	-4.86	-9.62		-5.12	-5.36	-5.50	-5.00
G8 NH	-6.98	-6.89	-6.94	-6.89	-6.98	-8.94			-2.22	-8.15	-7.12
E9 NH	-5.29	-5.41	-5.81	-6.06	-5.93	-6.47	-6.19	-6.00		-7.83	-6.13
L10 NH	-5.72	-5.94	-5.73	-5.93	-5.96	-6.72	-8.42	-5.90			
Q11 NH	-5.21	-5.94	-5.60	-5.49	-5.20	-6.31	-6.16	-6.40		-5.70	

Q11 $CONH_2$	-6.83	-6.40	-6.59	-6.81	-3.86	-6.65	-6.34	-6.57	-4.86	-6.21	
R12 NH	-4.69	-4.77	-4.84				-4.92	-6.66	-5.81		-5.89
R12 $CONH_2$	-4.02	-4.36	-4.64	-3.88	-1.67	-3.54	-3.20	-4.21	-3.22	-4.30	-4.98

TABLE 3

^1H NMR Chemical Shifts of Motilin 1-12 and Analogs in DMSO-d_6 at 30 °C

Residue	Chemical Shift (ppm)[a]				
	NH	αH	βH	γH	Others
Motilin 1-12					
Phe-1	8.17	4.16	2.88,3.01		ArH 7.17,7.40
Val-2	8.74	4.34	1.97	0.93,1.02	
Pro-3		4.51	2.04	1.83	δH 3.61
Ile-4	7.86	4.15	2.00	1.06	γH' 0.76
					δH 0.76
Phe-5	8.05	4.69	2.81,3.04		ArH 6.86,7.31
Thr-6	7.90	4.28	4.01	0.99	
Tyr-7	7.93	4.48	2.75,2.97		ArH 6.70,7.08
Gly-8	8.30	3.77			
Glu-9	8.03	4.39	1.78,1.91	2.28	
Leu-10	8.03	4.31	1.58	1.49	δH 0.86,0.93
Gln-11	8.06	4.20	1.91	2.13	εNH 6.90
Arg-12	7.84	4.18	1.69	1.48	δH 3.12
					εNH 7.55
NH_2	n.o.				

Ψ[CH_2NH]$^{1-2}$motilin 1-12

Phe-1	8.16	3.47	2.81,3.08		ArH 7.13,7.37
					CH_2 3.59
Val-2	7.89	4.16	1.66	0.75,1.00	
Pro-3		4.42	1.96	1.76	δH 3.48,3.58
Ile-4	7.91	4.31	1.81	1.35	γH' 0.87
					δH 0.85
Phe-5	8.10	4.71	2.83,3.05		ArH 6.86,7.27
Thr-6	7.87	4.29	4.03	1.00	
Tyr-7	7.92	4.49	2.78,2.99		ArH 6.74,7.05
Gly-8	8.28	3.82			
Glu-9	8.06	4.35	1.79,1.94	2.29	
Leu-10	8.01	4.31	1.65	1.50	δH 0.89,0.96
Gln-11	8.05	4.23	1.80,1.94	2.15	εNH 7.93
Arg-12	7.82	4.24	1.71	1.50,1.57	δH 3.13
					εNH 7.56
NH_2	n.o.				

Ψ[CH_2NH]$^{4-5}$motilin 1-12

Phe-1	7.99	3.86	2.89		ArH 7.12,7.36
Val-2	8.73	4.35	1.99	0.95	
Pro-3		4.33	2.05	1.85	δH 3.63
Ile-4	8.62	4.15	n.o.	n.o.	γH' n.o.
					δH 0.92
					CH_2 n.o.
Phe-5	8.17	4.16	3.02		ArH 6.80,7.24
Thr-6	8.56	4.37	3.96	1.06	
Tyr-7	8.08	4.48	2.74,2.98		ArH 6.70,7.08
Gly-8	8.28	3.78			
Glu-9	8.03	4.32	1.78,1.94	2.28	
Leu-10	8.04	4.22	1.50	1.64	δH 0.89,0.92
Gln-11	8.06	4.22	1.81	2.15	εNH n.o.
Arg-12	7.82	4.21	1.71	1.51,1.57	δH 3.12
					εNH 7.54
NH_2	7.36				

[a] n.o.; not observed.

TABLE 4

$^3J_{NH\alpha H}$ Coupling Constants for Motilin 1-12 and Reduced Analogs

Residue	Ψ[CH_2NH] Position		
	None	1-2	4-5
Phe-1	5.82		
Val-2	6.79	6.84	7.25
Pro-3			
Ile-4	6.77	6.56	
Phe-5	6.48	6.34	6.73
Thr-6	7.32	7.20	
Tyr-7	6.15	6.25	7.24
Gly-8	7.11	7.01	
Glu-9	8.16	8.36	6.70
Leu-10	7.25	7.15	8.07
Gln-11	6.08	6.18	7.51
Arg-12	6.32	6.41	6.23

FIGURE LEGENDS

FIGURE 1

NH-aliphatic region of the NOESY spectra (600.13 MHz) of motilin 1-12 recorded in DMSO-d_6 at 300K using a mixing time of 250 ms.

FIGURE 2

Superposition of specific regions of the molecular models of (a) motilin 1-12 (Phe^1-Ile^4), (b) $\Psi[CH_2NH]^{4-5}$motilin 1-12 (Phe^1-Val^2) and (c) $\Psi[CH_2NH]^{4-5}$motilin 1-12 (Gln^{11}-NH_2^t) where the formation of CO···HN hydrogen bonding occurs.

FIGURE 3

Superposition of ten low energy structures of (a) motilin 1-12, (b) $\Psi[CH_2NH]^{1-2}$motilin 1-12 and (c) $\Psi[CH_2NH]^{4-5}$motilin 1-12 calculated by molecular modeling.

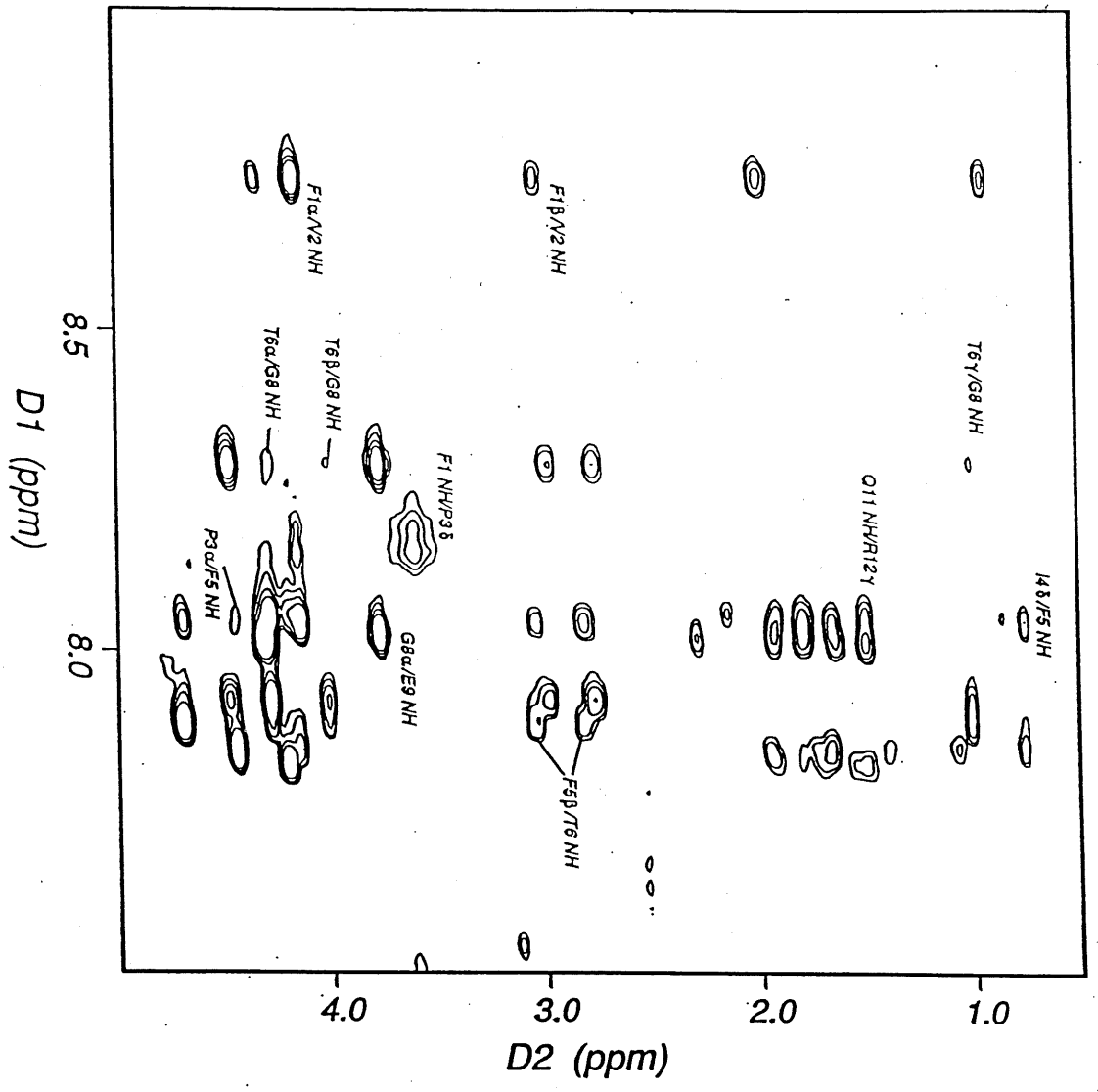
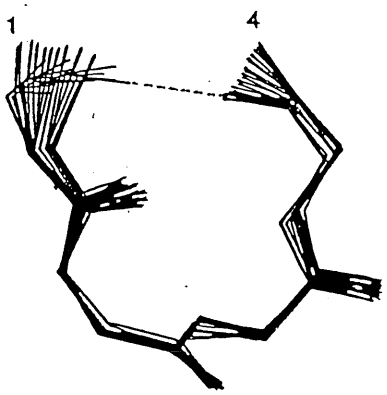
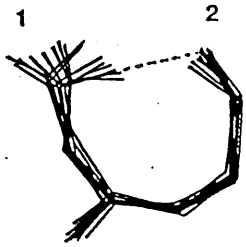

F1α/V2 NH
F1β/V2 NH
T6γ/G8 NH
T6α/G8 NH
T6β/G8 NH
F1 NH/P3δ
Q11 NH/R12γ
I4δ/F5 NH
P3α/F5 NH
G8α/E9 NH
F5β/T6 NH
8.5
8.0
D1 (ppm)
4.0
3.0
2.0
1.0
D2 (ppm)

Fig. 1

A



B



C

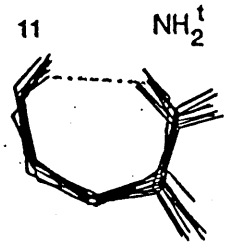


Fig. 2

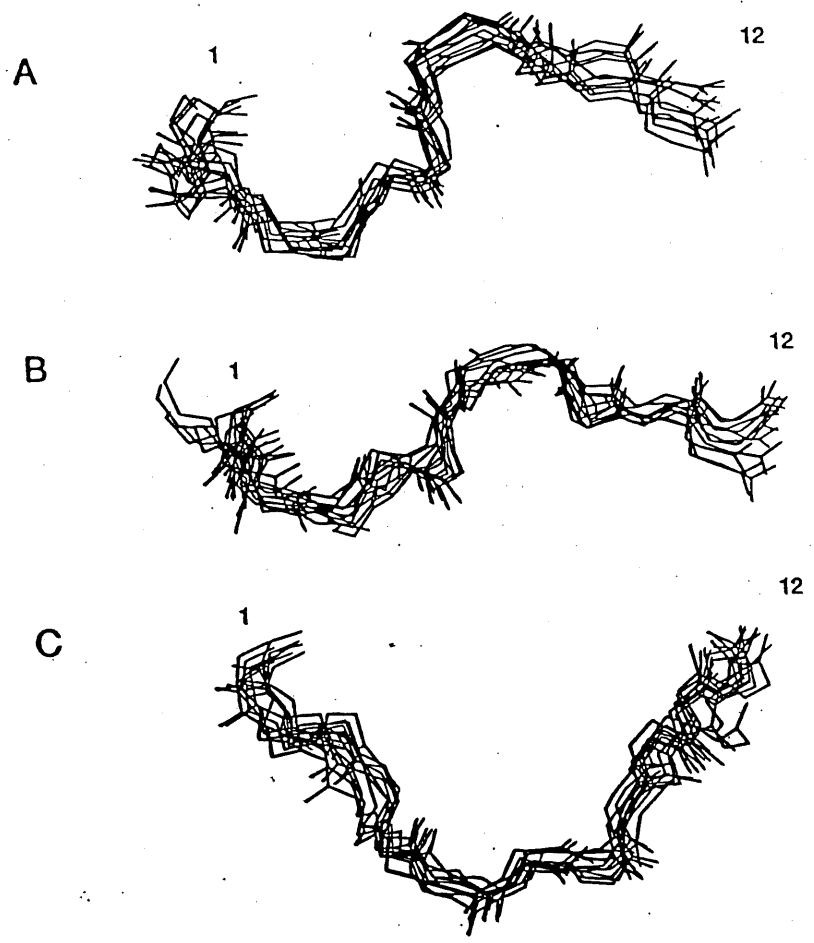


Fig. 3

CHAPITRE IV

Structural Differences between the Free and Bound States of the DNA-Bisintercalating Peptide YSPTSPSY

Abdesslem Khiat, and Yvan Boulanger*

INRS-Santé, Université du Québec, 245 boul. Hymus, Pointe-Claire, Québec, Canada H9R 1G6

Journal of medicinal chemistry, sous presse, Juin 1996

INTRODUCTION

Le peptide YSPTSPSY, un intercalateur de l'ADN, appartient à une famille de médicaments anticancéreux qui comprend la triostine A et l'échinomycin. Dans le cas de ces deux dernières molécules, la structure tridimensionnelle de leur complexe avec un fragment d'ADN a été résolue par RMN et cristallographie, par contre ces données structurales ne sont pas disponibles pour le complexe YSPTSPSY-ADN. Le but de ce travail est d'essayer d'explorer l'espace conformationnel accessible au peptide et au complex par une analyse multiconformationnelle. Une stratégie basée sur une recherche conformationnelle aléatoire avec minimisation d'énergie a été appliquée pour générer les populations de conformères qui caractérisent YSPTSPSY. L' analyse par les méthodes statistiques a permis de déterminer l'existence de quatre classes de conformères qui contiennent des tournants β et / ou γ. En utilisant une procédure d'arrimage basée sur la méthode de Monte-Carlo, le peptide YSPTSPSY été arrimé dans un fragment d'ADN en double hélice avec la séquence $d(GACGTC)_2$. Le peptide s'intercale dans le petit sillon de l'ADN entre les paires de bases CG, d'une manière similaire à celle observée dans le complexe de la triostine A avec l'ADN. En se liant à l'ADN, la structure du segment C-terminal est modifiée en tournant β de type I. Cinq liaisons hydrogène intermoléculaires sont observées mais les interactions de type van der Waals constituent le facteur majeur de stabilisation du complexe.

J'ai effectué le travail de recherche et de rédaction de la première partie de cet article au département de chimie de l'Université de Sherbrooke, la deuxième partie (recherche et rédaction) à l'INRS-Santé.

Abstract

Experimental evidences suggest that YSPTSPSY, a DNA-bisintercalating peptide, can adopt non-random conformations in aqueous solution. Strategies based on random conformational search and energy minimizations have been applied to generate populations of conformers characterizing YSPTSPSY. Subsequent analysis based on statistical methods and clustering allowed to determine the existence of four classes of conformers containing β- and/or γ-turns. Employing a Monte-Carlo based docking procedure, the YSPTSPSY peptide was docked in a DNA double helical fragment with the sequence $[d(GACGTC)]_2$. The peptide binds on the minor groove of DNA stacking the central CG base pairs, in a manner similar to that observed in complexes of triostin A with DNA. Upon binding, the structure of the C-terminal segment is modified into a type I β-turn. Five intermolecular hydrogen bonds are observed but the van der Waals interactions constitute the major stabilization factor for the complex.

Introduction

The heptad repeat unit SPTSPSY of RNA polymerase II has been proposed to bind to transcription factors (1). Extending this heptad unit with one tyrosine at the N-terminus yields the octapeptide YSPTSPSY (peptide I) which has been shown to bind to DNA (2) as a member of a peptide family containing two terminal planar rings (3-5). This peptide contains two overlapping Ser-Pro-X-X sequences often found in gene regulatory proteins (6). Experimental evidence based on fluorescence measurements (2) show that peptide I binds to DNA by the intercalation of tyrosine residues. In the model proposed by Suzuki (2), both Ser-Pro-X-X sequences form type-I β-turns, one being interlocked to the other. The first β-turn is formed by Ser^2-Pro^3-Thr^4-Ser^5 with the existence of two intramolecular hydrogen bonds, one between the carbonyl oxygen of Ser^2 and the amine hydrogen of Ser^5 and the other between the β-hydroxyl oxygen of Ser^2 and the amine hydrogen of Thr^4. The second β-turn is formed by the Ser^5-Pro^6-Ser^7-Tyr^8 segment and involves hydrogen bonding between the carbonyl oxygen of Ser^5 and the amine hydrogen of Tyr^8. An additional hydrogen bond was also postulated between the side chain oxygen of Ser^5 and the amine hydrogen of Ser^7. Formation of these two β-turns serves to orient the two terminal tyrosine rings in adequate position for spanning two CG base pairs of DNA (2). As demonstrated by NMR analysis (7), there is good evidence for the presence of type I and type II β-turns in the Ser^2-Pro^3-Thr^4-Ser^5 segment, but the Ser^5-Pro^6-Ser^7-Tyr^8 segment is less structured. In this study, we used molecular modeling calculations to demonstrate the existence not only of the structure proposed by Suzuki (2) but also of structures comprising one β-turn and one γ-turn and structures comprising two γ-turns. These structures are very stable with the terminal tyrosine residues positioned favorably for interaction with the minor groove of DNA.

Triostin A is a bicyclic antibiotic with a repeated sequence of four amino acids $(D\text{-}Ser\text{-}Ala\text{-}N\text{-}MeCys\text{-}N\text{-}MeVal)_2$ containing two planar aromatic quinoxaline rings attached to the D-serine residues (3, 9). This molecule is C-shaped with a central

region containing the cyclic depsipeptide with the two quinoxaline rings perpendicular to the peptide backbone and parallel to each other. This arrangement orients the quinoxaline rings in favorable position to intercalate into DNA. The crystal and NMR structures of the triostin A-DNA complex (8-10) show that triostin A binds to the minor groove as bisintercalator around the CG base pairs. The alanine residues of the peptide are positioned so that their NH and carbonyl groups are located inside the C-shaped structure at the interface with the nucleic acid. These alanine residues stabilize the complex by forming four hydrogen bonds with the nucleic acid bases: two hydrogen bonds are formed between the amine hydrogen of the alanines and the N3 of the guanines and two hydrogen bonds are formed between the carbonyl oxygens of the alanines and the NH_2 hydrogens of the guanines. However, the major binding interaction between the drug and the nucleic acid involves a large number of van der Waals contacts.

In this study, the mode of bisintercalation of peptide I into DNA was investigated by docking calculations and was found to involve a rearrangement of the C-terminal structure. Details of the structures and interactions are presented and compared to the crystal and NMR structures of the triostin A-$[d(GACGTC)]_2$ complex.

Experimental Section

Random Conformational Search. The peptide sequence of YSPTSPSY (peptide I) was built using the standard residue library of the ECEPP (empirical conformation energy program for peptides) forcefield (11, 12). The PEPSEA (peptidic search) method (13) was used which, instead of searching a global minimum, generates the most representative peptide population. A total of 15,000 conformers was generated by randomly changing 24 torsion angles (Φ, Ψ, χ) of the backbone and the side chains. Dihedral angles were labelled according to the IUPAC-IUB Commission on Biochemical Nomenclature (14). The conformational energies of the

resulting structures were calculated using the ECEPP forcefield and energy minimized using the conjugate gradients algorithm (15) for 1000 iterations with a RMS gradient of 0.001 kcal.mol^{-1}.Å^{-1}. The validity of the population sample was assessed by the bell shape of the energy histogram. Individual analysis of the first 100 structures allowed a classification of the conformers into characteristic families. The calculations were performed on IBM RISC/6000 model 520 workstation (Laboratoire de modélisation moléculaire, Université de Sherbrooke, Canada). The SYBYL molecular modeling program (Tripos Associates Inc., St. Louis MO) was used for calculating and analyzing the molecular structures.

Docking Procedure. Docking studies were carried out on the Silicon Graphics workstation using the software modules INSIGHTII and AUTODOCK (16) from Biosym Technologies, San Diego CA. The Biosym consistent valence forcefield (CVFF) was used to build the peptide and the B-DNA structure, determined by NMR in the presence of triostin A (10), was obtained from the Protein Data Bank (Brookhaven National Laboratory, Upton NY). Peptide I was docked in the minor groove of the DNA fragment using the Monte Carlo docking procedure. This procedure uses a high temperature dynamics calculation. A key component of the procedure is the Monte Carlo-driven insertion of peptide conformers into the DNA structure based on the random positioning within the DNA molecule followed by a calculation of the interaction energy with the affinity grid. Displacing the peptide results either in an increased or a decreased affinity energy. When the affinity energy is increased, the positioning is rejected and a new random positioning is chosen. When the affinity energy is decreased, the positioning is accepted and displacement is pursued in the same direction until the interaction energy is practically constant. The resulting complex, which is often in a highly energetic state, is then energy minimized using the conjugate gradients algorithm until the RMS gradient is less than 0.001 kcal.mol^{-1}.Å^{-1}.

Results and Discussion

Conformational Domains of Peptide I. Conformations of peptide I generated by the random conformational search have been classified in accordance with the conventional letter-coded regions of the (Φ,Ψ) maps (17). The most stable conformers are distributed into four major families presented in Figure 1. Table 1 presents a list of the structured conformers among the first 100 calculated structures divided in families with their identification number, conformational code, type of β-turn, conformational energy, frequency of occurrence and presence of hydrogen bonds. All hydrogen bonds were shorter than 2.9 Å and their XH···Y angles larger than 100°.

The first family, which comprises the majority of structures, is characterized by a β-turn in the N-terminal segment Ser^2-Pro^3-Thr^4-Ser^5 (Table 1). Most conformers of this family show the presence of hydrogen bonding between the carbonyl oxygen of Ser^5 and the amine hydrogen of Ser^7, forming a γ-turn in the C-terminal segment. An important number of conformers of this family are stabilized with a third hydrogen bond between the side chain hydroxyl oxygen of Ser^2 and the amine hydrogen of Thr^4. Conformers of the second family are characterized by two hydrogen bonds, one between the carbonyl oxygen of Ser^2 and the amine hydrogen of Thr^4, the other between the carbonyl oxygen of Ser^5 and the amine hydrogen of Ser^7. These hydrogen bonds give rise to two γ-turns, one in each extremity of the peptide chain. The third family comprises conformers containing a β-turn in the C-terminal segment Ser^5-Pro^6-Ser^7-Tyr^8, along with the presence of hydrogen bonding between the carbonyl oxygen of Ser^5 and the amine hydrogen of Tyr^8. Most conformers of this family display a second hydrogen bond between the carbonyl oxygen of Ser^2 and the amine hydrogen of Thr^4. Some conformers of this family are stabilized with a third hydrogen bond between the side chain oxygen of Ser^5 and the amine hydrogen of Ser^7. Finally, the fourth family is characterized by two β-turns located, respectively, in the N- and C-terminal segments. Figure 2 shows the superimposition of structures belonging to the four families. Several structures belonging to none of the four families remained. These

presented other hydrogen bonding patterns or were extended structures.

The largest number of structures belongs to family 1 which includes the most stable conformers, energetically and thermodynamically (frequency of occurrence). On the other hand, conformers with two β-turns (family 4) are the least stable energetically and the least abundant (Table 1). These results also indicate that the C-terminal portion of peptide I is less structured than the N-terminal portion, in agreement with the NMR results (7).

Analysis of the structures in each family allows a characterization of the geometrical parameters, especially the distances between the β-carbons and hydroxyl oxygens of the terminal tyrosine residues. In most conformers, the aromatic rings of both tyrosines are located on the same side of the molecule and their Cβ-Cβ and hydroxyl O-O distances are comparable to the reported values of 11.1 Å and 10.2 Å, respectively (2). Superimposition of one structure of each family with the structure proposed by Suzuki (2) is illustrated in Figure 3 and demonstrates that the presence of two β-turns is not necessary to position the aromatic rings on the same side of the molecule with Tyr^1-Tyr^8 distances suitable for intercalation into DNA. In fact, conformers of families 1, 2 and 3 containing γ-turn structures present their terminal tyrosine rings in positions very similar to the Suzuki model (2).

Our approach with multiconformational analysis allows us to underline two important points. First, clear evidence is obtained that the structure of unbound peptide I is not a random coil but that the conformational space is populated with structures in which the Ser-Pro-X-X segments form mostly type III β-turns and occasionally type I or type II β-turns (the structures of type I and type III β-turns are very similar (18)). Second, comparing with Suzuki's model (2), we have been able to show (Figure 3) that structures with a N-terminal β-turn and a C-terminal γ-turn (family 1), with a N-terminal γ-turn and a C-terminal β-turn (family 3) or with two γ-turns (family 2) are very stable presenting in adequate position the terminal tyrosine rings for intercalation into DNA.

Structure of the Peptide I-DNA Complex. Figure 4 displays the peptide I-$[d(GACGTC)]_2$ complex as viewed in the major groove (Figure 4A) and sideways (Figure 4B). As can be seen, peptide I binds to the minor groove with the two tyrosine rings stacking on each side of the CG base pairs. Figure 5 illustrates the stacking arrangement of the nucleic acid bases and tyrosine rings as viewed from the top of the helix. Different degrees of ring overlap are observed. A closer view of the interaction between peptide I and the nucleic acid fragment is shown in Figure 6. Torsion angles of the nucleic acid and peptide I backbone are given in Table 2. The geometry of the oligonucleotide, which had been determined by NMR in the presence of triostin A, is only slightly changed in the presence of peptide I.

A major stabilizing factor for the peptide I-DNA complex is provided by the existence of a large number of van der Waals interactions, the calculated energy of interaction being -94.03 kcal/mol (Figure 7). A total of 36 van der Waals distances shorter than 3.5 Å are present between the atoms of peptide I and those of the DNA fragment. An important stabilizing factor is the aromatic ring stacking which is maximal when the inter-ring distance is 3.4 Å. The Tyr^1 ring largely overlaps with the G^4 ring at a distance of 3.40 Å and partially with the A^8 ring at a distance of 3.50 Å. The aromatic ring of Tyr^8 overlaps strongly with the G^{10} ring at a distance of 3.52 Å and with the A^2 ring but at the longer distance of 4.00 Å (Figure 5). In addition, van der Waals interactions exist between Ser^2, Ser^5 and the central CG base pairs. Five intermolecular hydrogen bonds are formed: between the amine hydrogen of Ser^2 and the NH_2 nitrogen of G^4, between the carbonyl oxygen of Ser^2 and the NH_2 hydrogens of G^{10}, between the side chain hydroxyl hydrogen of Ser^2 and the NH_2 nitrogen of G^4, between the side chain hydroxyl hydrogen of Ser^5 and the NH_2 nitrogen of G^8 as well as between the carbonyl oxygen of Ser^7 and the N3 of A^2. This suggests that Ser^2 and Ser^5 are the key residues for the binding of peptide I to the central CG base pairs.

Comparison of Peptide I-DNA and Triostin A-DNA Complexes. Similarities are found between the peptide I-DNA complex and the triostin A-DNA complex. In both complexes, strong van der Waals contacts and four hydrogen bonds between the central CG base pairs and the serine or alanine residues, for peptide I or triostin A, respectively. The stacking interactions between the tyrosine rings and the base pair rings in the peptide I-DNA complex are similar to those observed between the six-membered ring of quinoxaline in triostin A and the base pair rings in the triostin A-DNA complex. In both complexes, multiple van der Waals interactions constitute the major component of the stabilization (8-10).

Superimposition of the intercalated structures of peptide I and triostin A (Figure 8) shows strong similarities between the two molecules. Both tyrosine rings in the case of peptide I and both quinoxaline rings in the case of triostin A are perpendicular to the main chain, parallel together and positioned favorably for intercalation. Both peptides are C-shaped and Thr^4 in peptide I seems to play the same structural role as the disulfide bridge in triostin A. In the peptide I structure, both terminal tyrosine residues are diagonally opposed to Thr^4 whereas Ser^2 and Ser^5 are pointing toward the nucleic acid and positioned to play a crucial role in complex formation. A similar situation is observed for triostin A where the serine residues are diagonally opposed to the disulfide bridge and the two alanine residues are pointing toward the nucleic acid in a favorable position for bisintercalation.

Conformation of Peptide I in the Complex. The lowest energy conformation of peptide I, belonging to family 1 (conformation 1 in Table 1) has been docked into the DNA fragment. This conformation was chosen because, according to our results, the majority of conformers belongs to family 1. It contains both a N-terminal β-turn (residues Ser^2-Pro^3-Thr^4-Ser^5) and a C-terminal γ-turn (residues Ser^5-Pro^6-Ser^7). The docking results indicate that in the complex of peptide I with the central CG base pairs, the major stabilization factor is due to the interaction between the N-terminal part of peptide I and the CG base pairs. These interactions are only possible if the peptide

adopts a conformation in which the N-terminal segment is ideally distanced from the CG base pairs, which occurs only when peptide I is folded to form a β-turn in its N-terminal part. As illustrated in Figure 9, it is evident that conformers of family 2 or family 3 with a N-terminal γ-turn cannot bind to DNA, even though their tyrosine rings can intercalate. Atoms from the N-terminal segment remain far from the CG base pairs and, therefore, interactions essential to the formation of the complex cannot occur. In our attempts to dock a conformer of family 2 containing two γ-turns, the peptide was always expelled from the DNA binding site.

Following docking and energy minimization of the complex, the structure of the family 1 conformer rearranged itself inside the complex. Formation of an intermolecular hydrogen bond between the carbonyl oxygen of Ser^2 and the NH_2 nitrogen of G provoked the rupture of the intramolecular hydrogen bond between the carbonyl oxygen of Ser^2 and the amine hydrogen of Ser^5. This leads to favorable interactions between the N-terminal segment of peptide I and the central CG base pairs. Concomitantly, the structure of the C-terminal segment rearranged itself in a more folded conformation to allow a better insertion of the Tyr^8 ring into the A^2-T^{11} and C^3-G^{10} base pairs. Analysis of the torsion angles of peptide I (Table 2) in the final complex structure, indicates that the values of the Φ and Ψ angles of Pro^3 and Thr^4 are very similar to the angles characterizing a type I β-turn, except for the Ψ angle of Thr^4 which deviates by approximately 40° from the ideal value. The major structural change, however, occurs in the C-terminal segment of the peptide which changes from a γ-turn to a type I β-turn better defined than in the N-terminal segment. This new geometry of peptide I complexed to the DNA fragment is relatively similar to the structure proposed by Suzuki (2) solely on the basis of a comparison between peptide I and triostin A. It is particularly interesting to note the similarity of the interatomic distances between the terminal aromatic rings. The distances between the β-carbons and the hydroxyl oxygens of Tyr^1 and Tyr^8 were, respectively, of 11.1 Å and 10.2 Å in Suzuki's model and 11.07 Å and 10.08 Å, respectively, in our model. Comparison of the peptide I

conformations before and after the docking procedure suggests that the conformation of the peptide is modified due to bisintercalation and that the structures of peptide I free and bound to DNA are different.

References

1. Sigler, P.B. Transcriptional Activation. Acid Blobs and Negative Noodles. *Nature* **1988**, *333*, 210-212.
2. Suzuki, M. The Heptad Repeat in the Largest Subunit of RNA Polymerase II Binds by Intercalating to DNA. *Nature* **1990**, *344*, 562-565.
3. *Chemistry of Antitumour Agents*. Wilman, D.E., Ed.; Blackie & Son Limited: Glasgow 1990; pp 1-59; pp 403-436.
4. Wakelin, L.P.E. Polyfunctional DNA Intercalating Agents. *Med. Res. Revs.* **1986**, *6*, 275-339.
5. Waring, M.J.; Wakelin, P.G. Echinomycin: A Bifunctional Intercalating Antibiotic. *Nature* **1974**, *252*, 653-657.
6. Suzuki, M. SPXX, a Frequent Sequence Motif in Gene Regulatory Proteins. *J. Mol. Biol.* **1990**, *207*, 61-84.
7. Harding, M.M. NMR Studies of YSPTSPSY: Implications for the Design of DNA Bisintercalators. *J. Med. Chem.* **1992**, *35*, 4658-4664.
8. Wang, A.H.-J.; Ughetto, G.; Quigley, G.J.; Hakoshima, T.; van der Marel, G.A.; van Boom, J.H.; Rich, A. The Molecular Structure of a DNA-Triostin A Complex. *Science* **1986**, *225*, 1115-1121.
9. Ughetto, G.; Wang, A.H.-J.; Quigley, G.J.; van der Marel, G.A.; van Boom, J.H.; Rich, A. A Comparison of the Structure of Echinomycin and Triostin A Complexed to a DNA Fragment. *Nucleic Acids Res.* **1985**, *13*, 2305-2323.
10. Addess, K.J.; Feigon, J. Sequence Specificity of Quinoxaline Antibiotics. 1. Solution Structure of a 1:1 Complex Between Triostin A and $[d(GACGTC)]_2$ and Comparison with the Solution Structure of the [N-MeCys^3,N-MeCys^7]TANDEM-$[d(GATATC)]_2$ Complex. *Biochemistry* **1994**, *33*, 12386-12396.
11. Momany, F.A.; McGuire, R.F.; Burgess, A.W.; Scheraga, H.A. Energy Parameters in Polypeptides. VII. Geometric Parameters, Partial Atomic Charges, Nonbonded Interactions, Hydrogen Bond Interactions, and Intrinsic

Torsional Potentials for the Naturally Occurring Amino Acids. *J. Phys. Chem.* **1975**, *79*, 2361-2381.

12. Momany, F.A.; Carruthers, L.M.; McGuire, R.F.; Scheraga, H.A. Intermolecular Potentials from Crystal Data. III. Determination of Empirical Potential and Applications to the Packing Configurations and Lattice Energies in Crystals of Hydrocarbons, Carboxylic Acids, Amines, and Amides. *J. Phys. Chem.* **1975**, *78*, 1595-1620.
13. Hassani, C.A. *La Génération et l'Analyse des Populations de Structures Peptidiques en Vue d'Étudier les Hypersurfaces d'Énergie Conformationnelle en Mécanique Classique*, Ph.D. Thesis, Université de Sherbrooke, Sherbrooke, Canada, 1991.
14. IUPAC-IUB Commission on Biochemical Nomenclature. Abbreviations and Symbols for the Description of the Conformation of Polypeptide Chains. Tentative Rules (1969). *Biochemistry* **1970**, *9*, 3471-3479.
15. Powell, M.J.D. Restart Procedures for the Conjugate Gradient Method. *Math. Program.* **1977**, 241-254.
16. Goodsell, D.S.; Olson, A.J. Automated Docking of Substrates to Proteins by Simulated Annealing. *Proteins: Struct. Funct. Genet.* **1990**, *8*, 195-202.
17. Zimmerman, S.S.; Pottle, M.S.; Némethy, G.; Scheraga, H.A. Conformational Analysis of the 20 Naturally Occurring Amino Acid Residues Using ECEPP. *Macromolecules*, **1977**, *10*, 1-9.
18. Dyson, H.J.; Rance, M.; Houghten, R.A.; Lerner, R.A.; Wright, P. Folding of Immunogenic Peptide Fragments of Proteins in Water Solution I. Sequence Requirements for the Formation of a Reverse Turn. *J. Mol. Biol.* **1988**, *201*, 161-200.

Acknowledgments

The authors would like to acknowledge the assistance of Mr. Louis Senécal and Mr. Alain Larocque for computer programming and maintenance of the computer equipment. This work was supported by the Medical Research Council of Canada and the Institut National de la Recherche Scientifique.

Table 1. Characteristics of Structured Conformers of Peptide I among the First 100 Conformers Generated by Random Conformational Search

conformer number	Zimmerman code[a]	type of β-turn	energy kcal/mol	frequency of occurrence	H-bonds[b]
Family 1: N-terminal β-turn					
1	F*EAA	III	-41.01	1	3
2	FAAA	III	-39.39	2	2, 3
3	EAAE	III	-38.41	2	
4	EAAE	III	-36.56	1	
5	GDAA	III	-35.36	1	2, 3
6	GDAA	III	-34.09	1	
7	FA*AA	III	-33.79	1	
8	FAAE	III	-32.47	2	2, 3
10	EA*AA	III	-30.12	1	3
12	BEAA	III	-29.88	2	
16	H*FCA	I	-28.04	1	3
23	FECA	I	-25.01	1	2, 3
25	B*EAA	III	-24.22	2	
26	D*A*AA	III	-24.00	1	
31	FA*AA	III	-21.11	2	3
34	B*EAB	I	-19.92	2	3
36	FFAA	III	-18.56	1	2
39	B*EAA	III	-16.99	1	
41	FFAA	III	-16.23	1	3
42	GDAB	I	-15.67	1	2
45	BFAA	III	-13.82	2	3
47	D*FAA	III	-12.02	2	3

50	FA*CA*	II	-11.29	1	3
51	FFAA	III	-10.77	1	2
52	AA*AA	III	-10.10	1	2
53	BCAB	I	-9.31	1	3
54	B*EAD	III	-8.90	2	3
57	CFFA*	II	-8.20	1	
59	FF*AA	III	-6.56	2	3
60	FA*GG	III	-5.02	1	2
61	BA*AA	III	-4.41	1	3

Family 2[c]: two γ-turns

9	-31.41	1
14	-29.39	2
17	-27.98	1
22	-25.91	1
27	-22.7	1
33	-20.13	1
38	-17.36	1
40	-16.44	2
44	-14.21	1
46	-13.59	1
49	-11.57	1
58	-7.29	1
62	-3.07	2
63	-1.66	1

Family 3: C-terminal β-turn					
11	FAAA*	III	-30.07	2	1, 4
13	F*ARE	I	-29.65	1	4
15	ECD*E	II	-29.00	1	1, 4
18	DAAC	III	-27.88	1	1, 4
19	EAAG	III	-27.01	2	1, 4
20	DCGE	III	-26.73	1	1
21	FAAB	III	-26.20	2	1
24	EABA	I	-24.32	1	
28	A*AAF*	III	-22.22	1	4
30	A*CA*E	II	-21.84	1	
37	DCGC*	III	-17.57	1	1
43	FCA*E	II	-14.52	1	
Family 4: two β-turns					
29	FAADAAC	III, III	-22.14	2	2
32	DABEABA	I, I	-20.69	1	2, 4
35	FAAFCA*E	III, II	-19.60	1	

[a] Zimmerman codes of the amino acids in the β-turn segment (17): residues Ser^{2}-Pro^{3}-Thr^{4}-Ser^{5} in the N-terminal segment and Ser^{5}-Pro^{6}-Ser^{7}-Tyr^{8} in the C-terminal segment. [b] Hydrogen bonds correspond to: 1, between the carbonyl oxygen of Ser^{2} and the amine hydrogen of Thr^{4}; 2, between the β-hydroxyl oxygen of Ser^{2} and the amine hydrogen of Thr^{4}; 3, between the carbonyl oxygen of Ser^{5} and the amine hydrogen of Ser^{7}; 4, between the β-hydroxyl oxygen of Ser^{5} and the amine hydrogen of Ser^{7}. [c] All conformers of this family displayed two γ-turns, one formed by the hydrogen bonding between the carbonyl oxygen of Ser^{2} and the amine hydrogen of Thr^{4}, the second between the carbonyl oxygen of Ser^{5} and the amine hydrogen of Ser^{7}.

Table 2. Torsion Angles of Amino Acids and Nucleic Acid Bases of the Lowest Energy Peptide I - Oligonucleotide Complex

Oligonucleotide backbone base	torsion angle (°)[a]						
	α	β	γ	δ	ε	ζ	χ
A^2			56.80	127.81	-173.06	9.33	-138.21
C^3	92.61	173.92	-175.28	84.34	174.39	-86.42	-113.09
G^4	150.23	-126.08	-166.48	124.00	-78.58	-64.60	-118.30
T^5	-56.83	-38.12	-160.27	142.90	-162.63	-129.95	-137.86
A^8	-81.04	-178.78	56.15	136.80	-152.45	-69.64	-98.73
C^9	98.61	175.76	-167.08	80.73	176.15	-119.73	-109.09
G^{10}	136.15	-103.33	-174.45	124.40	-69.04	-168.41	-125.96
T^{11}	92.94	-129.78	-174.78	143.43			-132.07

Peptide I

residue	torsion angle (°)[a]					
	Φ	Ψ	ω	X_1	X_2	X_3
Tyr[1]		125.06	-176.13	-102.12	-117.00	-179.35
Ser[2]	-166.95	96.00	161.39	50.10	-6.00	
Pro[3]	**-58.02**	**-30.66**	163,83			
Thr[4]	**-80.62**	**41.16**	172.08	-59.95	170.47	-176.59
Ser[5]	-80.15	152.09	163.62	81.74	-128.37	
Pro[6]	**-64.36**	**-9.55**	166.77			
Ser[7]	**-92.30**	**-2.79**	177.84	65.77	-50.99	
Tyr[8]	-134.80	78.48		-82.14	-93.15	-169.33

[a] Torsion angles of the oligonucleotide are defined as P$^{\alpha}$ O5$^{\beta}$ C5$^{\gamma}$ C4$^{\delta}$ C3$^{\varepsilon}$ O3$^{\zeta}$ P and O1'-C1'$^{\chi}$N9-C4 in purines or O1'-C1'$^{\chi}$N1-C2 in pyrimidines. Torsion angles of peptide I follow the standard nomenclature (13).

Figure Legends

Figure 1. Chemical structures of the four families of peptide I obtained following random conformational search and energy minimization. The location of hydrogen bonds leading to the formation of β- or γ-turns existing in all structures and to the formation of possible γ-turns are given by continuous and dotted lines, respectively.

Figure 2. Superimposition of structures generated by random conformational search and energy minimization for (a) family 1, (b) family 2, (c) family 3 and (d) family 4 of peptide I.

Figure 3. Superimposition of peptide I structures (thick line) belonging to (a) family 1, (b) family 2, (c) family 3 and (d) family 4 with the peptide I structure proposed by Suzuki (narrow line) (2).

Figure 4. Structure of the complex formed by peptide I (thick line) intercalated into the $[d(GACGTC)]_2$ oligonucleotide (narrow line) following docking calculations (a) front view into the major groove and (b) side view. The terminal nucleic acid base pairs are not shown. Several intermolecular distances (in Å) are given in part (a).

Figure 5. Sections of the peptide I-$[d(GACGTC)]_2$ complex showing the overlap of aromatic rings from the oligonucleotide bases (narrow lines) and peptide I tyrosines (thick lines) at the level of the (a) T^5-A^8, (b) G^4-C^9, (c) C^3-G^{10} and (d) A^2-T^{11} base pairs.

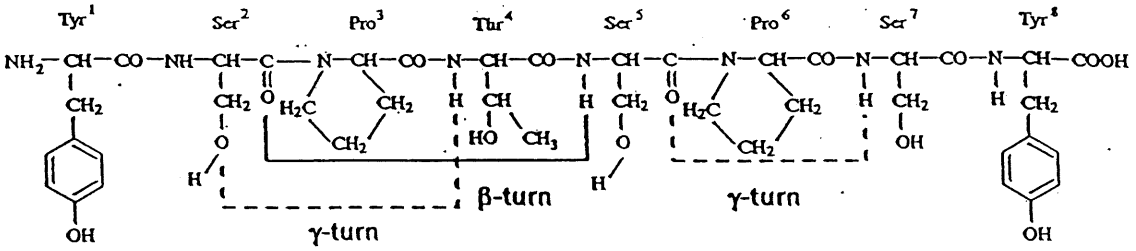
Figure 6. Section of the peptide I-$[d(GACGTC)]_2$ complex showing the intermolecular hydrogen bonds (thick dotted lines) and the distances between the aromatic rings of Tyr^1 and Tyr^8 in peptide I (narrow dotted lines). Distances are given in Å.

Figure 7. Graph showing the evolution of the energy of interaction of the peptide I-$[d(GACGTC)]_2$ complex during the docking and conjugate gradients energy minimization calculations.

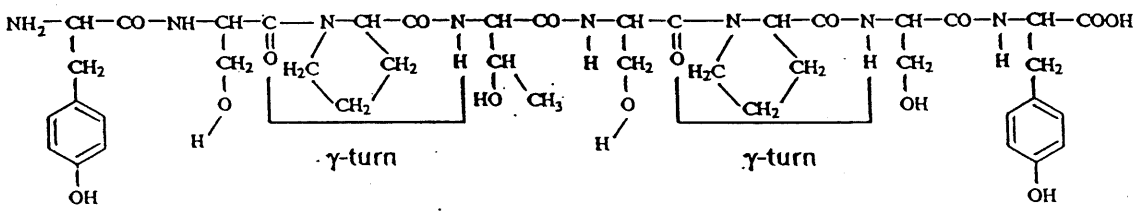
Figure 8. Comparison of the structures of peptide I (thick lines) and triostin A (narrow lines) (10) when bound to the same oligonucleotide. Structural similarities are evident, Thr^4 in peptide I overlapping with the disulfide bridge in triostin A.

Figure 9. Structure of the peptide I-$[d(GACGTC)]_2$ complex showing the difference between a peptide I conformer from family 2 (thick lines) and the energy minimized docked conformer (dotted lines). An important conformational difference is observed in the N-terminal segment which results in the impossibility of binding for a family 2 conformer with two γ-turns.

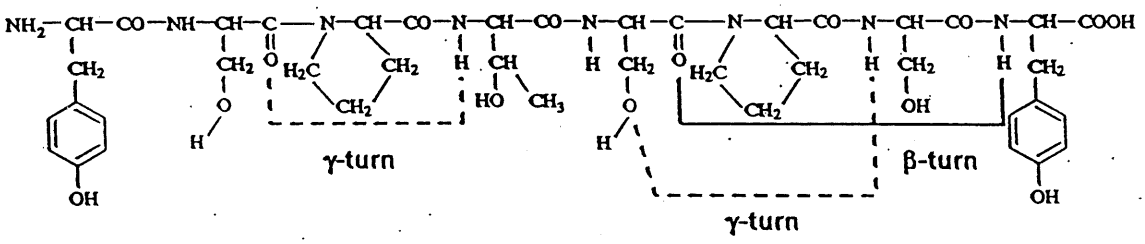
Family 1



Family 2



Family 3



Family 4

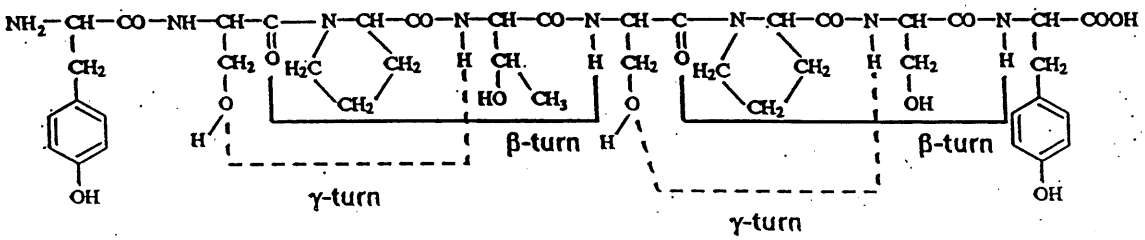


Fig. 1

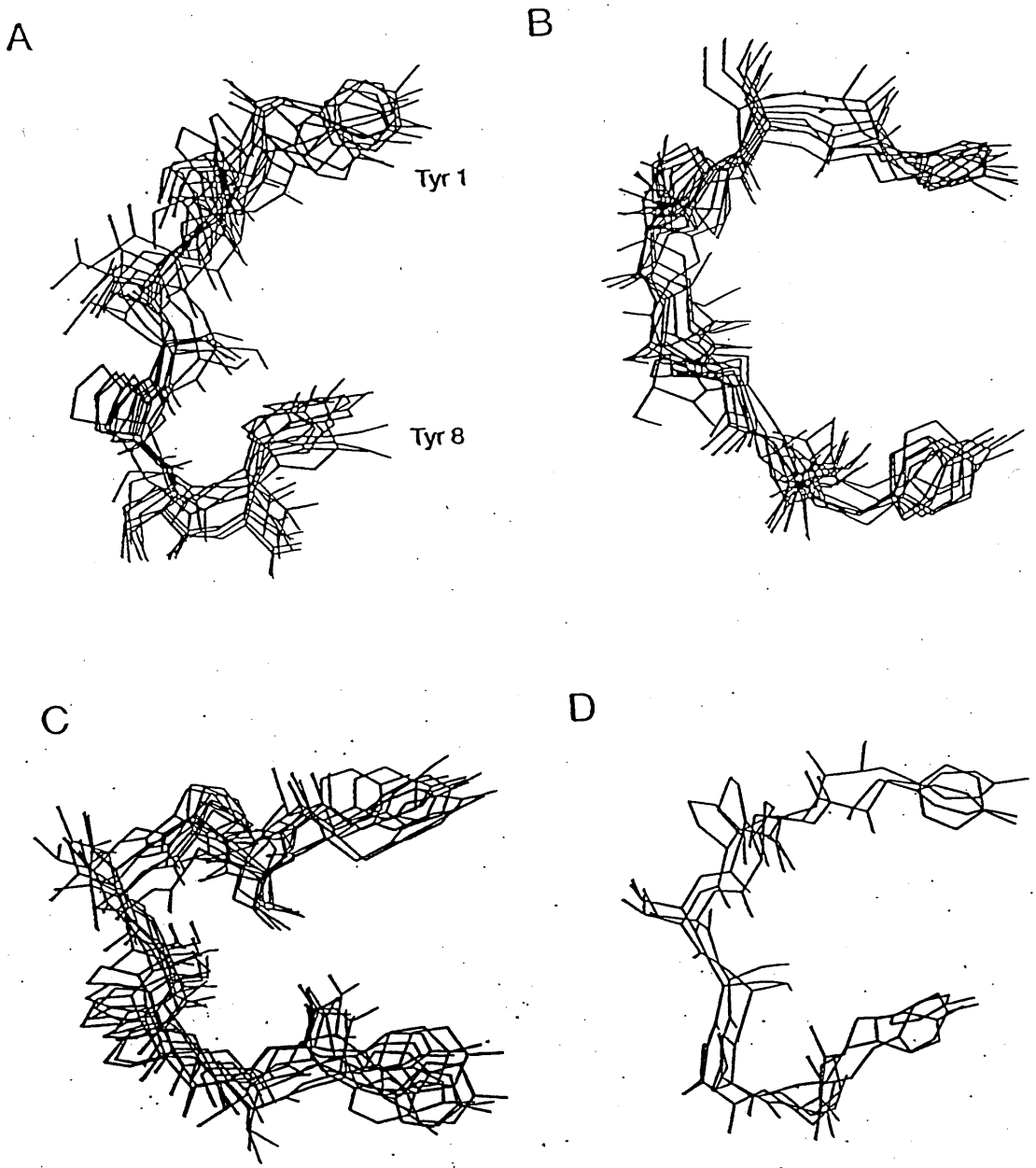


Fig. 2

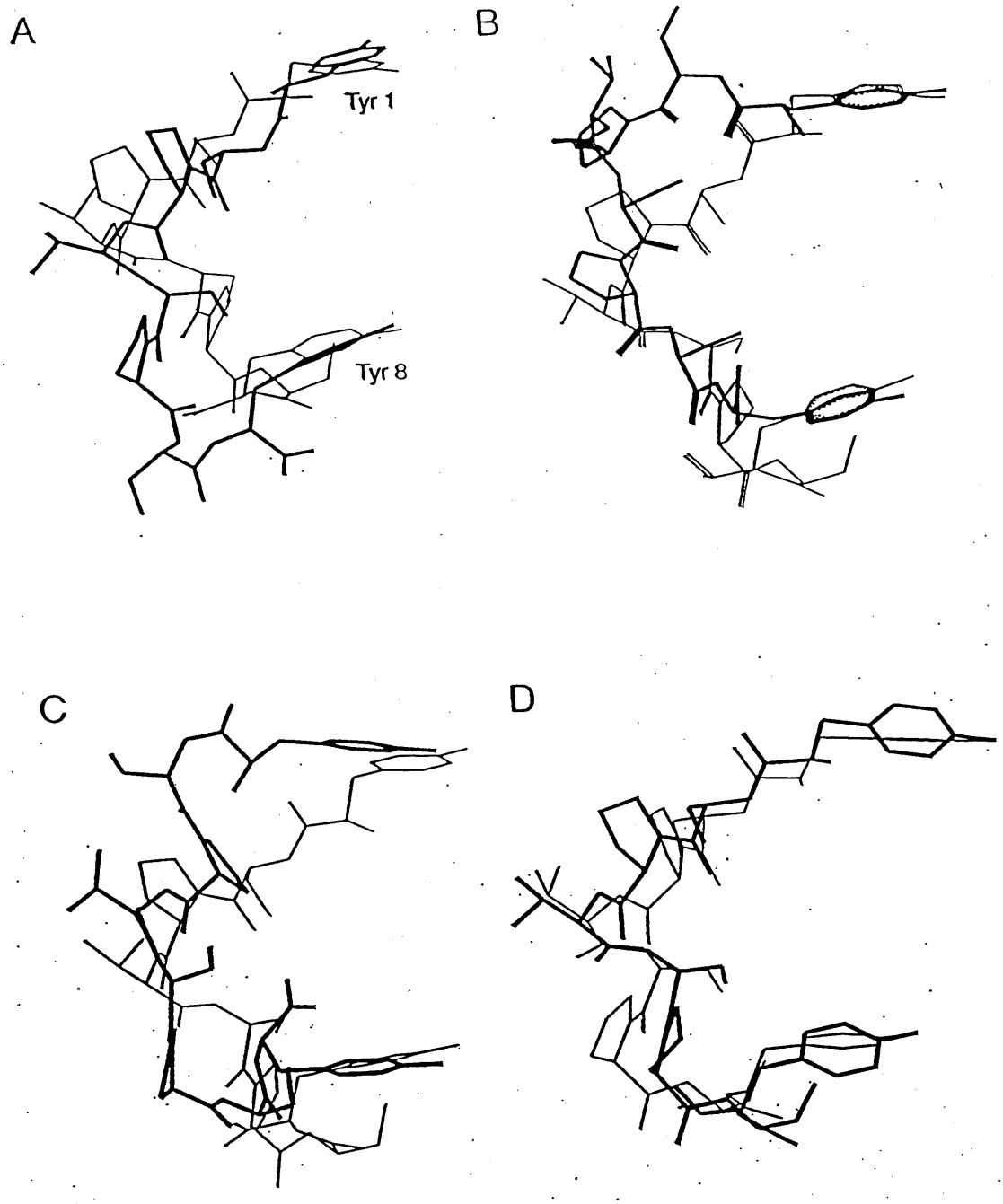


Fig. 3

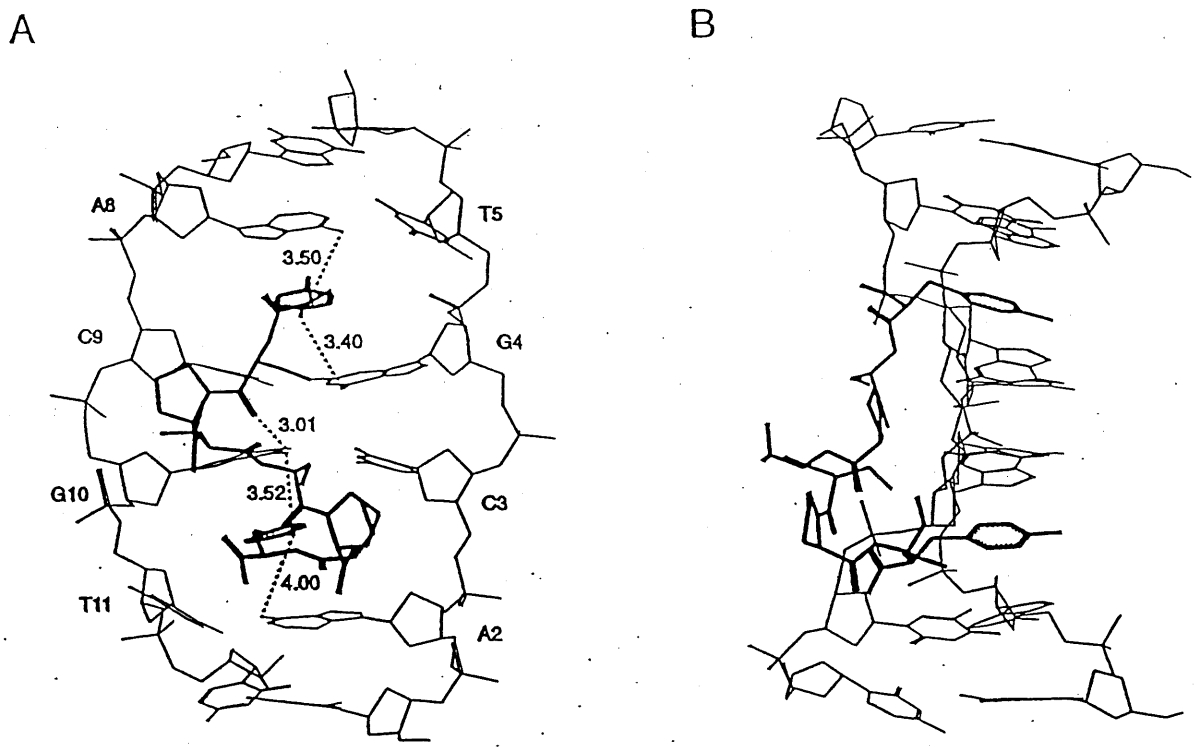


Fig. 4

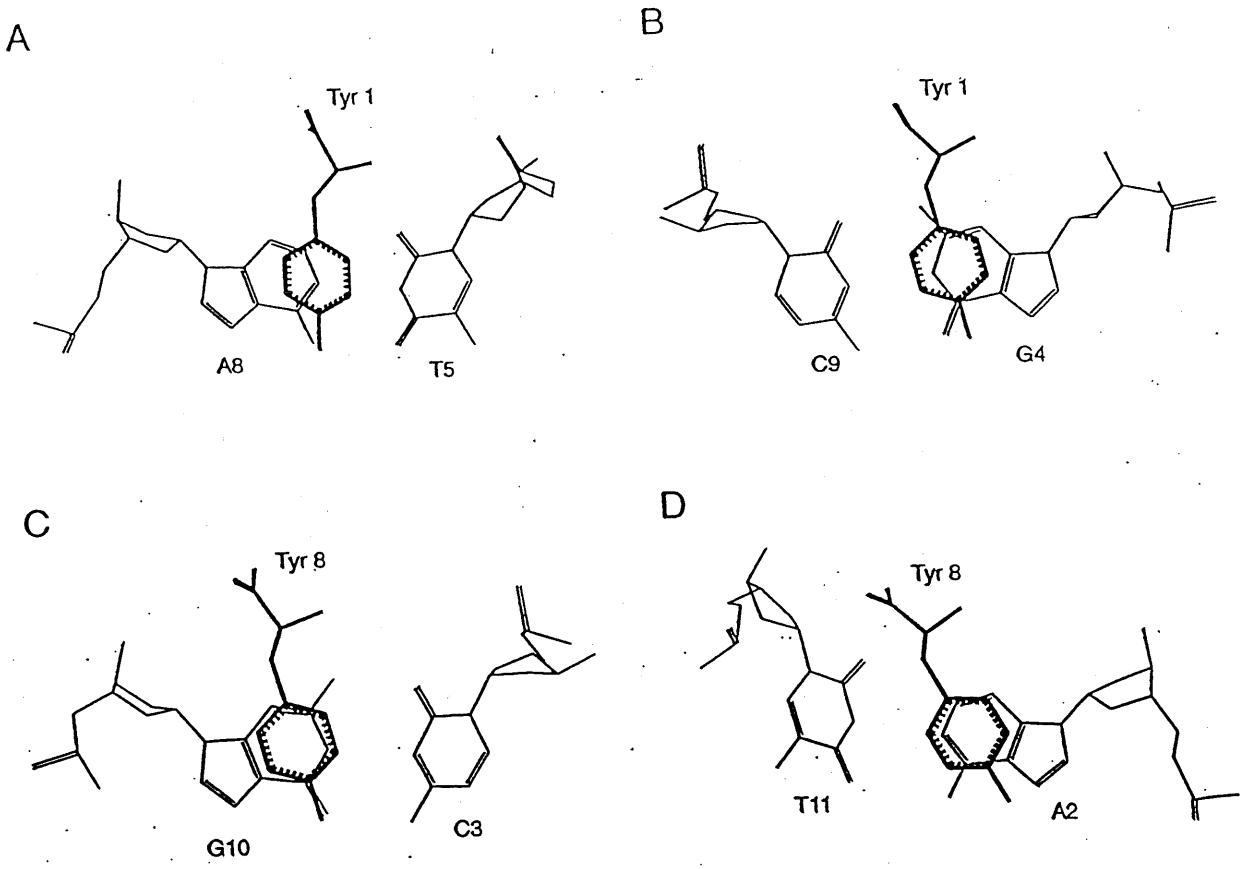


Fig. 5

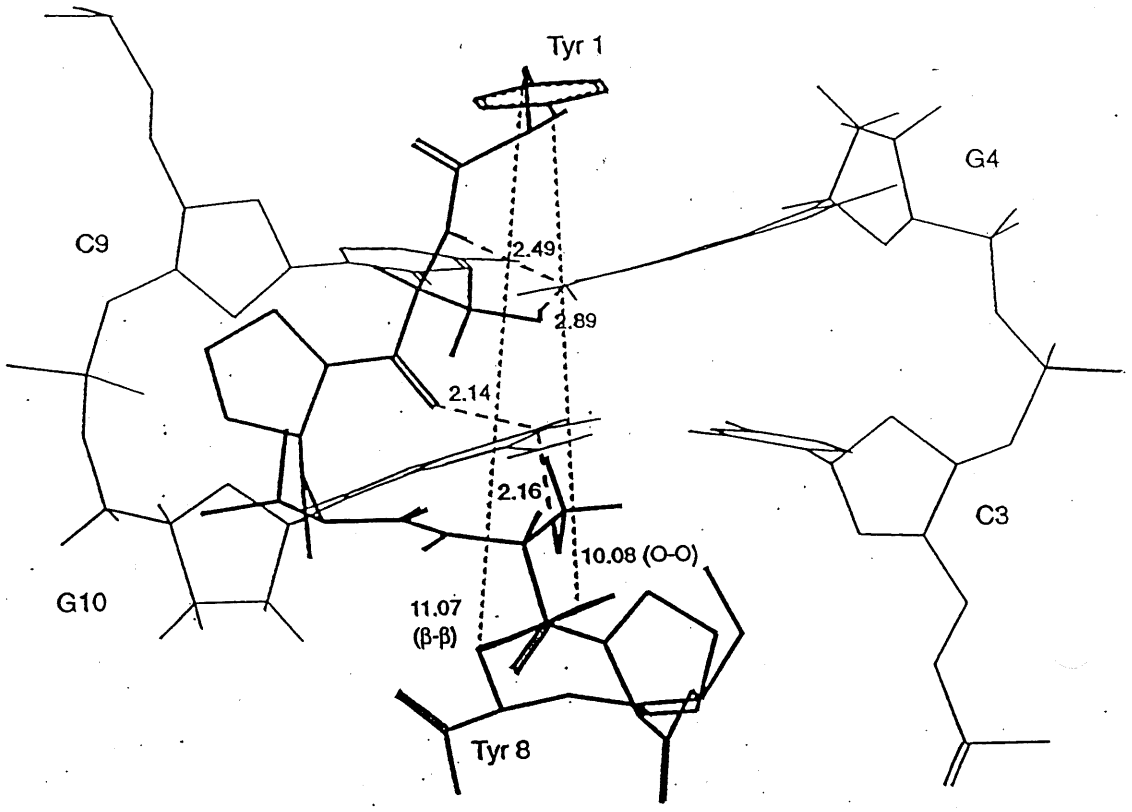


Fig. 6

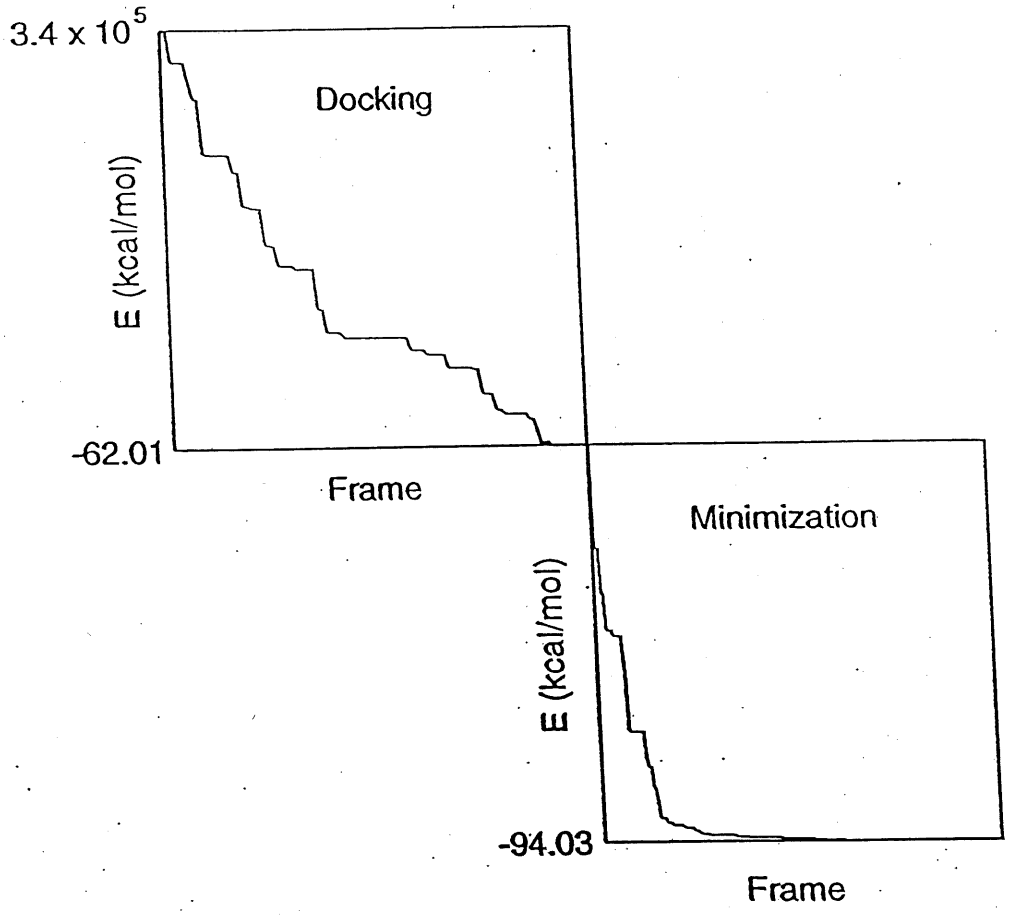

3.4 x 10^5
Docking
E (kcal/mol)
-62.01
Frame
Minimization
E (kcal/mol)
-94.03
Frame

Fig. 7

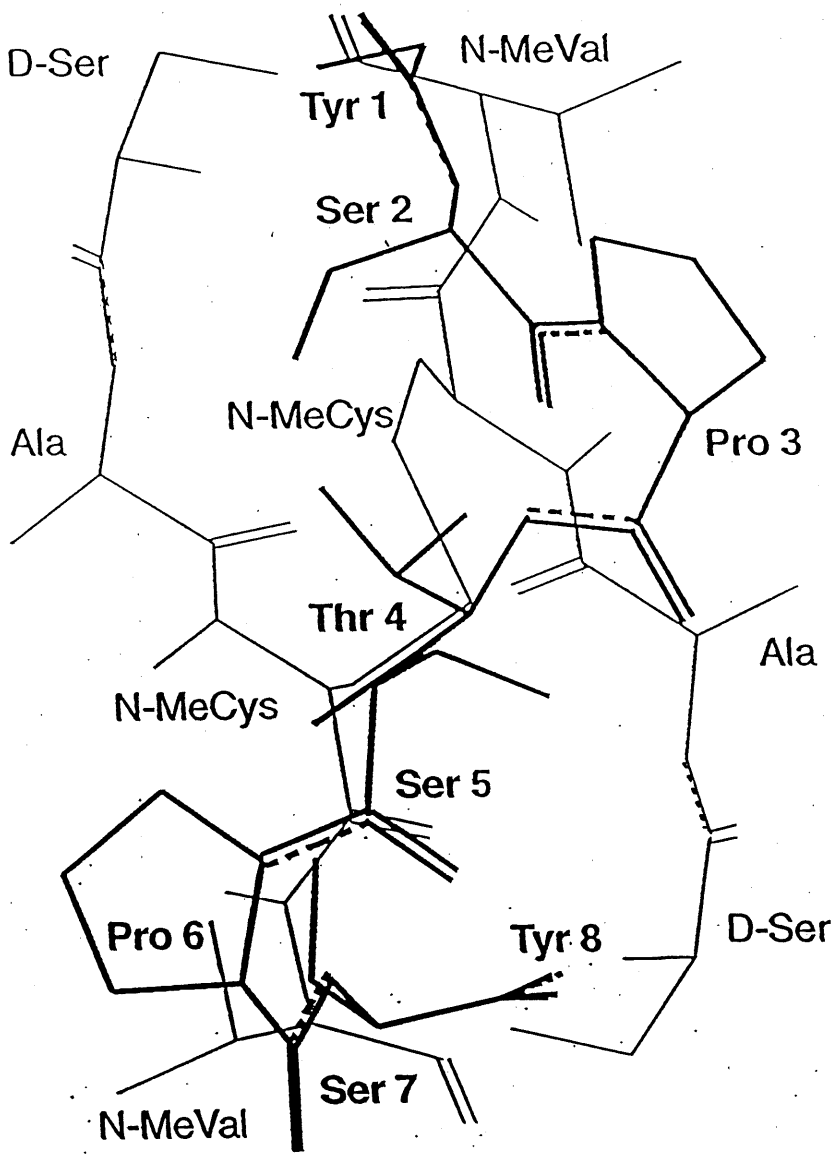

D-Ser
N-MeVal
Tyr 1
Ser 2
N-MeCys
Ala
Pro 3
Thr 4
Ala
N-MeCys
Ser 5
Pro 6
Tyr 8
D-Ser
Ser 7
N-MeVal

Fig. 8

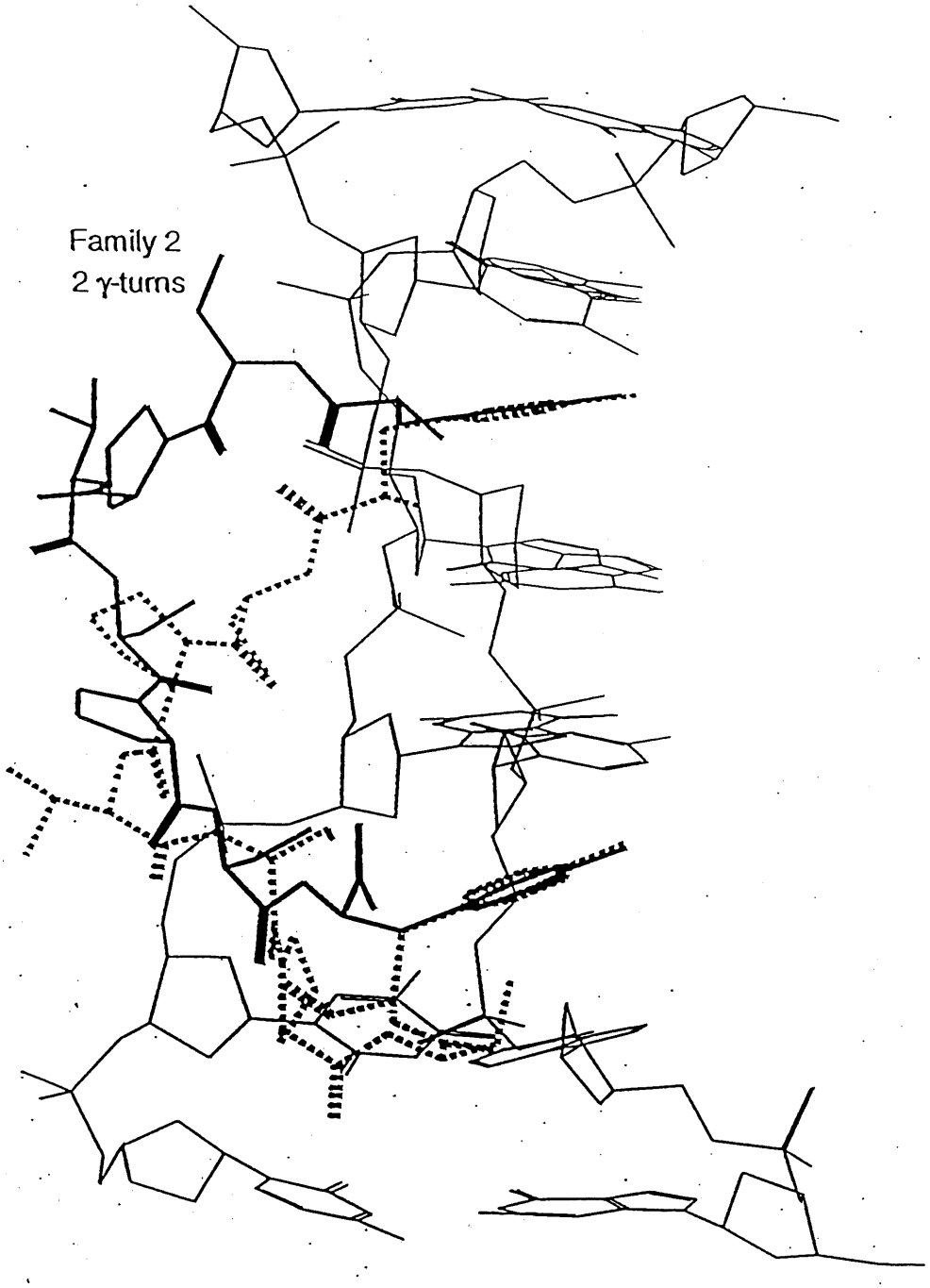


Fig. 9

CHAPITRE V

Structural comparison of alanine-substituted analogues of the calcitonin gene-related peptide 8-37: Importance of the C-terminal segment for antagonistic activity

Y. BOULANGER, A. KHIAT, A. LAROCQUE,
A. FOURNIER and S. ST-PIERRE

INRS-Santé, Université du Québec, 245 boul. Hymus,
Pointe-Claire, Québec, H9R 1G6, Canada

International Journal of Peptide and Protein Research, sous presse, Juin 1996

INTRODUCTION

Le remplacement de résidus spécifiques par un résidu alanine est une technique très utilisée pour augmenter le caractère agoniste ou antagoniste des peptides. Cette technique a été utilisée dans le cas du fragment du peptide alternatif du gène de la calcitonine 8-37 (hCGRP 8-37). Le remplacement par des résidus alanine en position 17 et 20 produit de bons antagonistes au récepteur CGRP1, mais un antagoniste médiocre en position 21. Le but de ce travail est d'essayer d'expliquer cette différence de comportement biologique en comparant les structures secondaires et tertiaires des quatre analogues hCGRP 8-37, (Ala^{17})-, (Ala ^{20})- et (Ala^{21})-hCGRP 8-37. Les structures en solution des quatre analogues ont été déterminées par RMN et modélisation moléculaire. La comparaison des déplacements chimiques et des variations de déplacement chimique des protons amide en fonction de la température montrent que ces données sont différentes pour (Ala^{17})-hCGRP 8-37 et (Ala ^{20})-hCGRP 8-37 relativement au hCGRP 8-37 sur les segments N-terminal et central mais pas sur le segment C-terminal (résidus 31-37). Dans le cas de (Ala ^{21})-hCGRP 8-37, les différences sont observées sur toute la chaine. Les modèles moléculaires calculés par distance-géométrie, recuit simulé et minimisation d'énergie montrent une homologie structurale entre (Ala^{17})-hCGRP 8-37, (Ala ^{20})-hCGRP 8-37 et hCGRP 8-37 pour le segment Asn^{31}-Phe^{37} ainsi qu'une liaison hydrogène entre Val ^{28} et Asn^{31}. Ces similarités structurales ne sont pas observées avec (Ala^{21})-hCGRP 8-37. La structure du segment C-terminal du hCGRP 8-37 apparait donc critique pour l'activité antagoniste.

Pour la réalisation de ce travail, j'ai acquis toutes les bases théoriques relatives à la modélisation moléculaire au département de chimie de l'Université de Sherbrooke, j'ai effectué le travail de recherche et de rédaction de la partie modélisation moléculaire et une partie de la RMN à l'INRS-Santé. Les mesures RMN ont été faites à l'Institut Armand-Frappier.

ABSTRACT

Replacement of specific residues of the antagonistic fragment human calcitonin gene-related peptide 8-37 (hCGRP 8-37) by alanine residues produces good antagonists to CGRP1 receptors when the replacement is made at positions 17 and 20 but a poor antagonist when the replacement is made at position 21. The solution structures of hCGRP 8-37 and of the three alanine analogues have been determined by two-dimensional ^{1}H NMR spectroscopy and molecular modeling. Following the complete assignment of the NMR spectra, a comparison of the chemical shifts and of the temperature dependence of the amide chemical shifts showed that these parameters differed for [Ala^{17}]-hCGRP 8-37 and [Ala^{20}]-hCGRP 8-37 relative to hCGRP 8-37 in the N-terminal and central segments but not in the C-terminal segment (residues 31-37). In the case of [Ala^{21}]-hCGRP 8-37, differences were observed all along the chain. Molecular modeling calculations were performed by distance-geometry, simulated annealing and energy minimization using NOE distance constraints. Molecular models showed a structural homology between [Ala^{17}]-hCGRP 8-37, [Ala^{20}]-hCGRP 8-37 and hCGRP 8-37 in the C-terminal segment Asn^{31}-Phe^{37} as well as hydrogen bonding between Val^{28} and Asn^{31}. These structural similarities are not observed with [Ala^{21}]-hCGRP 8-37. Therefore, the structure of the C-terminal segment of hCGRP 8-37 appears to be critical for antagonistic activity at CGRP1 receptors.

Key words: peptide structure, calcitonin gene-related peptide, nuclear magnetic resonance, molecular modeling, analogues.

Abbreviations: BOP, benzotriazol-1-yl-oxy-tris-(dimethylamino)phosphonium hexafluorophosphate; CGRP, calcitonin gene-related peptide; DMSO-d_6, dimethylsulfoxide-d_6; DQFCOSY, double-quantum-filtered correlation spectroscopy; hCGRP 8-37, fragment 8-37 of the human calcitonin gene-related peptide; hCGRPα, α-form ofthe calcitonin gene-related peptide; NOESY, nuclear Overhauser enhancement spectroscopy; RMSD, root-mean-square deviation; TOCSY, total correlation spectroscopy.

The calcitonin gene-related peptide (CGRP) is a 37-residue neuropeptide displaying multiple actions on neuronal tissue, skeletal muscle, the cardiovascular system and as an inflammatory mediator (1). Two types of receptors have been identified, CGRP1 and CGRP2 (1-3). The CGRP 8-37 fragment, devoid of the N-terminal disulfide-bridged cycle, has been found to be a potent CGRP antagonist at the CGRP1 receptor site (4-6). It has been reported to produce hypothermic and antinociceptive effects (7) and to inhibit oedema (8).

The solution structure of human CGRP 8-37 (hCGRP 8-37) has been studied by circular dichroism in 1,1,1,3,3,3-hexafluoro-2-propanol: water (9) and 2,2,2-trifluoroethanol: water (10) mixtures and by two-dimensional ^{1}H NMR spectroscopy in dimethylsulfoxide-d_6 (11). All studies display a significant loss of helix structure in hCGRP 8-37 relative to hCGRP. In the NMR study, hydrogen bonding between the amide carbonyl oxygen of Val^{28} and the amide NH proton of Asn^{31} was observed but the remaining structure is mostly disordered (11).

Replacement of specific amino acids by better helix-promoting residues, such as alanine (12), is a technique that has been applied to several peptides in order to enhance the helix character and hopefully produce analogues with better agonist or antagonist activity (13-15). Specific alanine replacement has been applied to hCGRP 8-37 and has produced analogues with increased or reduced antagonistic activity (Quirion, R., Dumont, Y., Fournier, A. and St-Pierre, S., unpublished). Table 1 presents the biological activities of hCGRP 8-37, [Ala^{17}]-hCGRP 8-37, [Ala^{20}]-hCGRP 8-37 and [Ala^{21}]-hCGRP 8-37 in two tissue preparations: guinea pig left atria which contains CGRP1 receptors and rat *vas deferens* which contains CGRP2 receptors. Compared to hCGRP 8-37, [Ala^{17}]-hCGRP 8-37 is approximately twice as potent antagonist to CGRP1 receptors, [Ala^{20}]-hCGRP 8-37 is equipotent and [Ala^{21}]-hCGRP 8-37 is three times less potent. hCGRP 8-37 and the three analogues are weak antagonists to CGRP2 receptors (Table 1).

In this study, we report the structural analysis of these three alanine analogues of hCGRP 8-37 (Table 1) performed by two-dimensional ^{1}H NMR spectroscopy and molecular modeling. Although the peptides do not display much regular structure, homology analysis allows to correlate the structures with the antagonistic properties towards the CGRP1 and

CGRP2 receptors. The C-terminal segment seems to be particularly important for antagonism against CGRP1 receptors.

EXPERIMENTAL PROCEDURES

Peptide synthesis

The peptides hCGRPα, hCGRP 8-37 and its alanine analogues were synthesized according to the solid-phase peptide synthesis method following a procedure described previously (5, 16). Briefly, the syntheses were performed with a homemade manual multireactor synthesizer, using a benzhydrylamine resin as solid support and benzotriazol-1-yl-oxy-tris-(dimethylamino)phosphonium hexafluorophosphate (BOP) as coupling agent. Side-chain protection of *α-tert*-butoxycarbonyl amino acids was as previously reported (5). Deprotection of side chains and cleavage of peptides from the resin were achieved by treatment with liquid hydrofluoric acid in the presence of *m*-cresol. After precipitation and washing with anhydrous diethyl ether, the crude peptides were extracted with pure trifluoroacetic acid, followed by evaporation. The oily material was dissolved in water and lyophilized. The formation of the disulfide bridge of the cyclic peptides was performed in a degassed 80% acetic acid solution (1 mg/ml peptide) to which an equal volume of iodine solution (13 mM), dissolved in the same solvent, was added. The final solution was incubated for 2.5 h with occasional shaking. Then, zinc dust was added until a discolored solution was obtained. After filtration, the solution was evaporated to 50 ml, and 250 ml 0.06% trifluoroacetic acid/H_2O solution was added before lyophilization. After alkaline treatment of peptides, the crude preparations were purified by preparative reverse-phase high-performance liquid chromatography. The purity of all peptides exceeded 98%. Purified peptides were characterized by analytical HPLC, capillary electrophoresis and amino acid analysis.

NMR spectroscopy

Synthetic [Ala^{17}]-hCGRP 8-37, [Ala^{20}]-hCGRP 8-37 and [Ala^{21}]-hCGRP 8-37 (7 mg) samples

were dried overnight under vacuum at 40°C and dissolved in 100% dimethylsulfoxide-d_6 (DMSO-d_6, 0.5 ml) to yield a concentration of 4.5 mM. The sample solutions were introduced in 5 mm NMR tubes and purged with argon gas to remove oxygen.

One- and two-dimensional ^{1}H NMR spectra were recorded at 500.13 MHz and at 298 K on a Bruker AMX2 500 spectrometer. Double-quantum-filtered correlation spectroscopy (DQFCOSY), total correlation spectroscopy (TOCSY, mixing times 60-90 ms) and nuclear Overhauser enhancement spectroscopy (NOESY, mixing times 100-300 ms) spectra were collected using the time proportional phase incrementation mode with 1024 data points and 256 t_1 increments and with 32, 16 and 64 transients, respectively. The temperature dependences of the NH chemical shifts were measured on TOCSY spectra acquired at 300, 305, 310, 315, 320, 325 and 330 K.

Processing of NMR spectra was performed on a Silicon Graphics Indigo R4000 XZ workstation using the program FELIX 2.30 (Biosym Technologies, Inc., San Diego CA). Prior to Fourier transformation, two-dimensional spectra were multiplied by a 30° shifted sinebell function. Data were zero-filled to yield a 2048 x 2048 matrix with a digital resolution of 3.0 Hz/point. Sequential assignments were performed using conventional methods (17). The observation of relatively sharp linewidths is inconsistent with dimer formation.

In order to obtain distance constraints, cross-peak intensities were estimated from the 200-ms NOESY spectra. NOE intensities were classified as strong, medium and weak corresponding to upperbound constraints of 3, 4 and 5 Å, respectively. When equivalent proton nuclei were present on a same group, a distance correction corresponding to the maximal distance between the pseudoatom and the protons was added to the constraint (17). Both intraresidue connectivities for protons separated by more than four bonds and interresidue connectivities were used as constraints for molecular modeling.

Molecular Modeling

Molecular modeling calculations were carried out on the Silicon Graphics workstation using softwares purchased from Biosym Technologies, Inc. (INSIGHT II, DISCOVER, DGII,

HOMOLOGY version 2.35). The extended structure of each peptide, built using the consistent valence forcefield, was first submitted to 500 iterations of conjugate gradient energy minimization with a RMS gradient of 0.001 Å followed by molecular dynamics simulation using 5 fs time steps to generate 10^5 conformations sampled at every 100 fs. Cluster analysis of these conformations on the basis of interatomic distances allowed to classify them into five families of conformations. The lowest energy conformation in each family was energy minimized for 1000 iterations using the steepest descent algorithm. The NMR constraints were then introduced in each of the five structures and the distance-geometry algorithm (DGII) was used to create an initial set of structures based on the experimental distance constraints. The structures were subsequently refined via 1000 iterations of conjugate gradients energy minimization. Twenty structures were generated for each of the five starting conformations. The quality of the final structures was analyzed on the basis of the number of NOE distance violations and backbone root-mean-square deviation (RMSD) values. A pairwise comparison of the peptide structures was performed by the HOMOLOGY program to determine the structurally conserved regions (18). The homology threshold criterion was seven residues with a RMSD value of less than 1 Å.

RESULTS AND DISCUSSION

Spectral Assignment

The assignments of the ^{1}H NMR signals of [Ala^{17}]-hCGRP 8-37, [Ala^{20}]-hCGRP 8-37 and [Ala^{21}]-hCGRP 8-37 were performed from the analysis of the DQFCOSY, TOCSY and NOESY spectra and comparing with the already reported spectra of hCGRP and hCGRP 8-37 (11). Most amino acid spin systems were identified from the connectivity patterns observed in DQFCOSY and TOCSY spectra. The aromatic signals of the histidine and phenylalanine residues were identified from the β-Ar NOESY connectivities. Sequential assignments were obtained from interresidual connectivities involving backbone and side chain protons. An unambiguous assignment could be obtained for all three analogues. Signals of lower intensity due to cis Pro^{29} could be observed in the region of residues 27-31.

The chemical shifts of hCGRP 8-37, [Ala^{17}]-hCGRP 8-37, [Ala^{20}]-hCGRP 8-37 and [Ala^{21}]-hCGRP 8-37 are very similar, as shown in Table 2. The main differences between the alanine analogues and CGRP 8-37 are observed in N-terminal segment. [Ala^{17}]-hCGRP 8-37 presents five significant chemical shift differences (more than 0.05 ppm) in the region of residues 12-24 whereas [Ala^{20}]-hCGRP 8-37 presents seven such differences in the region of residues 10-24. Two such differences are observed for [Ala^{21}]-hCGRP 8-37 at positions 22 and 35. Only [Ala^{21}]-hCGRP 8-37 shows a chemical shift difference in the C-terminal segment.

Temperature coefficients

The temperature dependences of the chemical shifts of the amide protons are given in Table 3. Significant differences (more than 0.5 ppm/K) in the values of the temperature coefficients between the alanine analogues and hCGRP 8-37 are observed mostly in the N-terminal segment before residue Lys^{24}. Most of these differences coincide with the chemical shift differences discussed above. Therefore, the majority of the differences in chemical shifts (Table 2) and in temperature coefficients (Table 3) between hCGRP 8-37 and the alanine analogues are observed for the N-terminal region up to Lys^{24}, suggesting that the structure

of this region is the most affected by the replacement. Moreover, both chemical shifts and temperature coefficients demonstrate that only [Ala^{21}]-hCGRP 8-37 presents differences for the C-terminal residues Lys^{35} and Ala^{36}.

Two-dimensional NMR

Following the analysis of the NOESY spectra, mostly intraresidue and short range interresidue connectivities were observed. For hCGRP 8-37, 174 intraresidue and 126 interresidue (93 sequential; 17 i,i+2; 12 i,i+3; 3 i,i+4; 1 long range) connectivities were identified (11). For [Ala^{17}]-hCGRP 8-37, 156 intraresidue and 122 interresidue (78 sequential; 27 i,i+2; 9 i,i+3; 1 i,i+4; 7 long range) connectivities could be identified. In the case of [Ala^{20}]-hCGRP 8-37, 163 intraresidue and 41 interresidue (35 sequential; 6 i,i+2) were observed. For [Ala^{21}]-hCGRP 8-37, 162 intraresidue and 83 interresidue (70 sequential; 10 i,i+2; 3 i,i+3) connectivities were identified. Diagonal plots for all four peptides are shown in Fig. 1. Since all spectra were recorded under identical conditions, the reduced number of interresidue connectivities for [Ala^{20}]-hCGRP 8-37 is most likely attributable to increased mobility, especially in the C-terminal segment. The absence of a large number of medium range connectivities is inconsistent with the presence of helical structure for all peptides. The significant presence of helical structure reported previously for hCGRP 8-37 in 1,1,1,3,3,3-hexafluoro-2-propanol: water (9) and 2,2,2-trifluoroethanol: water (10) mixtures on the basis of circular dichroism measurements is most likely attributable to the solvent difference.

Molecular modeling

Following the molecular modeling procedure described above, 51, 63, 38 and 49 of the 100 final structures converged onto an identical folding pattern for hCGRP 8-37, [Ala^{17}]-hCGRP 8-37, [Ala^{20}]-hCGRP 8-37 and [Ala^{21}]-hCGRP 8-37, respectively. The calculated backbone RMSD values were 1.52 Å, 1.60 Å, 1.93 Å and 1.81 Å for hCGRP 8-37, [Ala^{17}]-hCGRP 8-37, [Ala^{20}]-hCGRP 8-37 and [Ala^{21}]-hCGRP 8-37, respectively. For all four peptides, no NOE

distance violations exceeded 0.5 Å at the end of the molecular modeling procedure. Fig. 2 displays the superimposed backbones of ten converging structures of each peptide.

The lowest energy structures of each alanine analogue peptide was compared to low energy hCGRP 8-37 structures using the HOMOLOGY modeling program in order to identify the structurally conserved regions. These calculations indicate that the C-terminal segment Asn^{31}-Phe^{37} is structurally conserved in [Ala^{17}]-hCGRP 8-37 and [Ala^{20}]-hCGRP 8-37 with RMSD values of less than 0.75 Å (Fig. 3). Such is not the case for [Ala^{21}]-hCGRP 8-37. The structure of the N-terminal and central segments do not display any homology for all three analogue peptides. Hydrogen bonding was observed between the amide NH proton of Asn^{31} and the amide carbonyl oxygen of Val^{28} for hCGRP 8-37, [Ala^{17}]-hCGRP 8-37 and [Ala^{20}]-hCGRP 8-37. In the case of [Ala^{21}]-hCGRP 8-37, hydrogen bonding was observed between the amide proton of Asn^{31} and the β-hydroxyl oxygen of Thr^{30}.

Comparison of the structures of these four hCGRP 8-37 peptides demonstrates the importance of the C-terminal segment for biological activity at CGRP1 receptors. When the structure of this segment is conserved, as is the case for [Ala^{17}]-hCGRP 8-37 and [Ala^{20}]-hCGRP 8-37, the antagonistic activity of the peptide for CGRP1 receptors is conserved (Table 1). However, when this structure is altered, as is the case for [Ala^{21}]-hCGRP 8-37, the antagonistic activity for CGRP1 receptors is severely reduced. The existence of hydrogen bonding between the amide carbonyl oxygen of Val^{28} and the amide NH hydrogen of Asn^{31} appears to be an essential feature for activity. These results are in agreement with the fact that short C-terminal fragments such as hCGRP 23-37 (19) and [Tyr^{0}]-CGRP 28-37 (20) are CGRP1 antagonists. They can also be paralleled with the loss of activity observed for hCGRP in CGRP1 receptors when the terminal phenylalanine residue was removed or chemically modified (21-23). In the case of CGRP2 receptors, the antagonistic activity of all three alanine analogues was reduced relative to hCGRP 8-37, although it is very weak in all cases (Table 1). This could suggest that this activity is more associated with the N-terminal portion of the peptide, since all three analogues were affected in this region. Therefore, the present results clearly suggest that different portions of the hCGRP 8-37 are associated with the activities

at receptors CGRP1 ad CGRP2. This characteristic, if confirmed by structure-activity studies of other analogues, represents a key element for the design of new CGRP antagonists.

ACKNOWLEDGMENTS

The authors are grateful to Mr. Louis Senécal for computer maintenance and software developments. This work was supported by the Medical Research Council of Canada.

REFERENCES

1. Poyner, D.R. (1992) *Pharmac. Ther.* **56**, 23-51
2. Dennis, T.B., Fournier, A., St-Pierre, S. & Quirion, R. (1989) *J. Pharmac. Exp. Ther.* **251**, 718-725
3. Quirion, R., Van Rossum, D., Dumont, Y., St-Pierre, S. & Fournier, A. (1992) *Ann. N. Y. Acad. Sci.* **657**, 88-105
4. Chiba, T., Yamaguchi, A., Yamatani, T., Nakamura, A., Morishita, T., Inui, T., Fukase, M., Noda, T. & Fujita, T. (1989) *Am. J. Physiol.* **256**, E331-E335
5. Mimeault, M., Quirion, R., Dumont, Y., St-Pierre, S. & Fournier, A. (1992) *J. Med. Chem.* **35**, 2163-2168
6. Dennis, T., Fournier, A., Guard, S., St-Pierre, S. & Quirion, R. (1990) *J. Pharmac. Exp. Ther.* **254**, 123-128
7. Saxen, M.A., Smith, F.L. & Dunlow, L.D. (1994) *Life Sci.* **21**, 1665-1674
8. Escott, K.J. & Brain, S.D. (1993) *Br. J. Pharmacol.* **110**, 772-776
9. Mimeault, M., St-Pierre, S. & Fournier, A. (1993) *Eur. J. Biochem.* **213**, 927-934.
10. Hubbard, J.A.M., Martin, S.R., Chaplin, L.C., Bose, C., Kelly, S.M. & Price, N.C. (1991) *Biochem. J.* **275**, 785-788
11. Boulanger, Y., Khiat, A., Chen, Y., Senécal, L., Tu, Y., St-Pierre, S. & Fournier, A. (1995) *Peptide Res.* in press
12. Scholtz, J.M. & Baldwin, R.L. (1992) *Annu. Rev. Biophys. Biomol. Struct.* **21**, 95-118
13. St-Pierre, S., Lalonde, J.-M., Gendreau, M., Quirion, R., Regoli, D. & Rioux, F. (1981) *J. Med. Chem.* **24**, 370-376
14. Fournier, A., Couture, R., Magnan, J., Gendreau, M., Regoli, D. & St-Pierre, S. (1980) *Can. J. Biochem.* **58**, 272-280
15. St-Pierre, S., Gaudreau, P., Drouin, J.N., Regoli, D. & Lemaire, S. (1979) *Can. J. Biochem.* **57**, 1084-1089

16. Forest, M., Martel, J.C., St-Pierre, S. & Quirion, R. (1990) *J. Med. Chem.* **33**, 1615-1619
17. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York
18. Needleman, S.B. & Wunsch, C.D. (1970) *J. Mol. Biol.* **48**, 443-453
19. Rovero, P., Giuliani, S. & Maggi, C.A. (1992) *Peptides* **13**, 1025-1027
20. Chakder, S. & Rattan, S. (1990) *J. Pharmac. Exp. Ther.* **253**, 200-206
21. O'Connell, J.P., Kelly, S.M., Raleigh, D.P., Hubbard, J.A.M., Price, N.C., Dobson, C.M. and Smith, B.J. (1993) *Biochem. J.* **291**, 205-210
22. Thiebaud, D., Akatsu, T., Yamashita, T., Suda, T., Noda, T., Martin, R.E., Fletcher, A.E. & Martin, T.J. (1991) *J. Bone Miner. Res.* **6**, 1137-1142
23. Zaidi, M., Brain, S.D., Tippins, J.R., Di Marzo, V., Moonga, B.S., Chambers, T.J., Morris, H.R. & MacIntyre, I. (1990) *Biochem. J.* **269**, 775-780

TABLE 1

Antagonistic activities of hCGRP 8-37 and alanine analogues in two biological preparations

Peptide	EC_{50} (nM)[a]	
	Guinea pig left atria (CGRP1 receptors)	Rat *vas deferens* (CGRP2 receptors)
hCGRPα	10.4±0.7	1.5±0.1
hCGRP 8-37	99 ± 13	6.3 ± 0.5
[Ala^{17}]-hCGRP 8-37	206 ± 37	4.5 ± 0.8
[Ala^{20}]-hCGRP 8-37	101 ± 22	1.3 ± 0.2
[Ala^{21}]-hCGRP 8-37	30 ± 8	2.3 ± 0.5

[a] In the case of hCGRP 8-37 and analogues, EC_{50} is the concentration of hCGRPα necessary to produce 50% of the maximal response induced in the presence of 1 μM analogue peptides. This value is the mean ± standard deviation of values obtained for 5-15 tissue preparations.

TABLE 2

^1H Chemical shifts of hCGRP 8-37 and alanine analogues in DMSO-d_6

Residue	Chemical shifts (ppm)							
	hCGRP 8-37		[Ala^{17}]hCGRP 8-37		[Ala^{20}]hCGRP 8-37		[Ala^{21}]hCGRP 8-37	
	NH	Hα	NH	Hα	NH	Hα	NH	Hα
Val 8	8.04	3.75		3.77	3.76	8.05	3.76	
Thr 9	8.37	4.28	8.39	4.28	8.41	4.28	8.41	4.28
His 10	8.24	4.65	8.24	4.65	8.14	4.60	8.27	4.66
Arg 11	8.14	4.32	8.17	4.25	8.19	4.23	8.18	4.23
Leu 12	7.91	4.08	7.90	4.27	7.90	4.28	7.91	4.30
Ala 13	7.98	4.24	7.99	4.27	8.05	4.28	8.00	4.22
Gly 14	8.02	3.68	8.05	3.66	8.05	3.67	8.05	3.71
Leu 15	8.08	4.24	8.14	4.30	8.08	4.32	8.12	4.29
Leu 16	7.99	4.30	7.97	4.25	7.99	4.30	8.02	4.30
Ser 17	7.89	4.31	7.88	4.24 (Ala)	7.95	4.31	7.98	4.25
Arg 18	7.99	4.23	7.94	4.30	7.99	4.30	8.02	4.30
Ser 19	7.95	4.29	7.90	4.28	7.82	4.27	7.84	4.27
Gly 20	8.03	3.64	8.15	3.74	7.95	4.23 (Ala)	8.11	3.71
Gly 21	8.11	3.75	8.07	3.71	8.08	3.70	7.96	4.35 (Ala)
Val 22	8.12	4.19	7.84	4.24	7.73	4.24	7.72	4.15
Val 23	7.82	4.12	7.87	4.11	7.87	4.12	7.90	4.14
Lys 24	7.88	4.19	7.88	4.30	7.90	4.28	7.93	4.24
Asn 25	8.06	4.52	8.07	4.51	8.07	4.54	8.07	4.51
Asn 26	8.10	4.45	8.12	4.48	8.12	4.47	8.12	4.45
Phe 27	7.91	4.45	7.90	4.43	7.92	4.43	7.93	4.45

Val 28	7.87	4.23	7.88	4.27	7.89	4.23	7.88	4.23
Pro 29		4.48		4.48		4.48		4.48
Thr 30	7.88	4.19	7.91	4.20	7.91	4.19	7.91	4.19
Asn 31	8.00	4.60	8.01	4.62	8.02	4.61	8.02	4.61
Val 32	7.70	4.10	7.72	4.10	7.72	4.10	7.72	4.10
Gly 33	8.18	3.74	8.18	3.74	8.19	3.74	8.20	3.73
Ser 34	7.82	4.28	7.93	4.30	7.91	4.28	7.91	4.30
Lys 35	8.07	4.25	8.09	4.20	8.09	4.20	8.10	4.18
Ala 36	7.88	4.14	7.90	4.13	7.89	4.13	7.90	4.14
Phe 37	7.68	4.35	7.70	4.36	7.70	4.36	7.70	4.34

TABLE 3

Temperature coefficients of hCGRP 8-37 and alanine analogues

Residue	Temperature coefficient (x 10^3 ppm/K)			
	hCGRP 8-37	[Ala^{17}]hCGRP 8-37	[Ala^{20}]hCGRP 8-37	[Ala^{21}]hCGRP 8-37
Val 8				-4.36
Thr 9	-4.01	-3.66	-4.00	-3.64
His 10	-5.72	-4.85	-4.00	-4.71
Arg 11	-3.80	-3.78	-4.00	-3.64
Leu 12	-4.58	-4.36	-4.36	-4.21
Ala 13	-5.58	-5.29	-5.00	-5.71
Gly 14	-4.01	-4.00	-4.00	-3.78
Leu 15	-2.58	-2.79	-4.36	-3.29
Leu 16	-5.69	-5.04	-5.07	-5.07
Ser 17	-4.83	-4.86 (Ala)	-4.43	-4.71
Arg 18	-4.15	-5.00	-4.14	-4.21
Ser 19	-6.01	-4.79	-5.14	-5.21
Gly 20	-5.37	-5.00	-4.86 (Ala)	-4.42
Gly 21	-4.22	-3.64	-3.57	-4.21 (Ala)
Val 22	-5.55	-4.71	-4.43	-4.57
Val 23	-5.80	-5.79	-6.00	-5.90
Lys 24	-4.30	-5.64	-4.21	-4.14
Asn 25	-5.01	-4.57	-4.57	-4.79
Asn 26	-4.62	-4.40	-4.64	-4.42
Phe 27	-2.94	-3.00	-3.00	-3.00
Val 28	-5.47	-5.57	-5.21	-5.14
Thr 30	-7.22	-6.14	-6.07	-5.90

Asn 31	-2.47	-2.71	-2.64	-2.79
Val 32	-3.01	-3.29	-3.36	-3.14
Gly 33	-4.08	-4.21	-4.00	-4.00
Ser 34	-5.30	-4.79	-4.57	-5.00
Lys 35	-5.90	-5.43	-5.57	-5.10
Ala 36	-4.26	-4.00	-4.00	-3.71
Phe 37	-4.44	-4.43	-4.57	-4.42

FIGURES

FIGURE 1

Diagonal plot of interresidual NOE connectivities observed for (a) hCGRP 8-37, (b) [Ala^{17}]-hCGRP 8-37, (c) [Ala^{20}]-hCGRP 8-37 and (d) [Ala^{21}]-hCGRP 8-37 in DMSO-d_6. A filled square indicates that at least one backbone-backbone connectivity was observed between the two residues, a shaded square indicates that at least one backbone-sidechain connectivity was observed and an open square indicates that only sidechain-sidechain connectivities were observed.

FIGURE 2

Superposition of ten low energy structures for (a) hCGRP 8-37, (b) [Ala^{17}]-hCGRP 8-37, (c) [Ala^{20}]-hCGRP 8-37 and (d) [Ala^{21}]-hCGRP 8-37 calculated by molecular modeling using NOE constraints.

FIGURE 3

Superposition of the lowest energy structure for hCGRP 8-37, [Ala^{17}]-hCGRP 8-37 and [Ala^{20}]-hCGRP 8-37 showing the homology of the C-terminal segment (residues Asn^{31}-Phe^{37}). No homology was found with [Ala^{21}]-hCGRP 8-37.

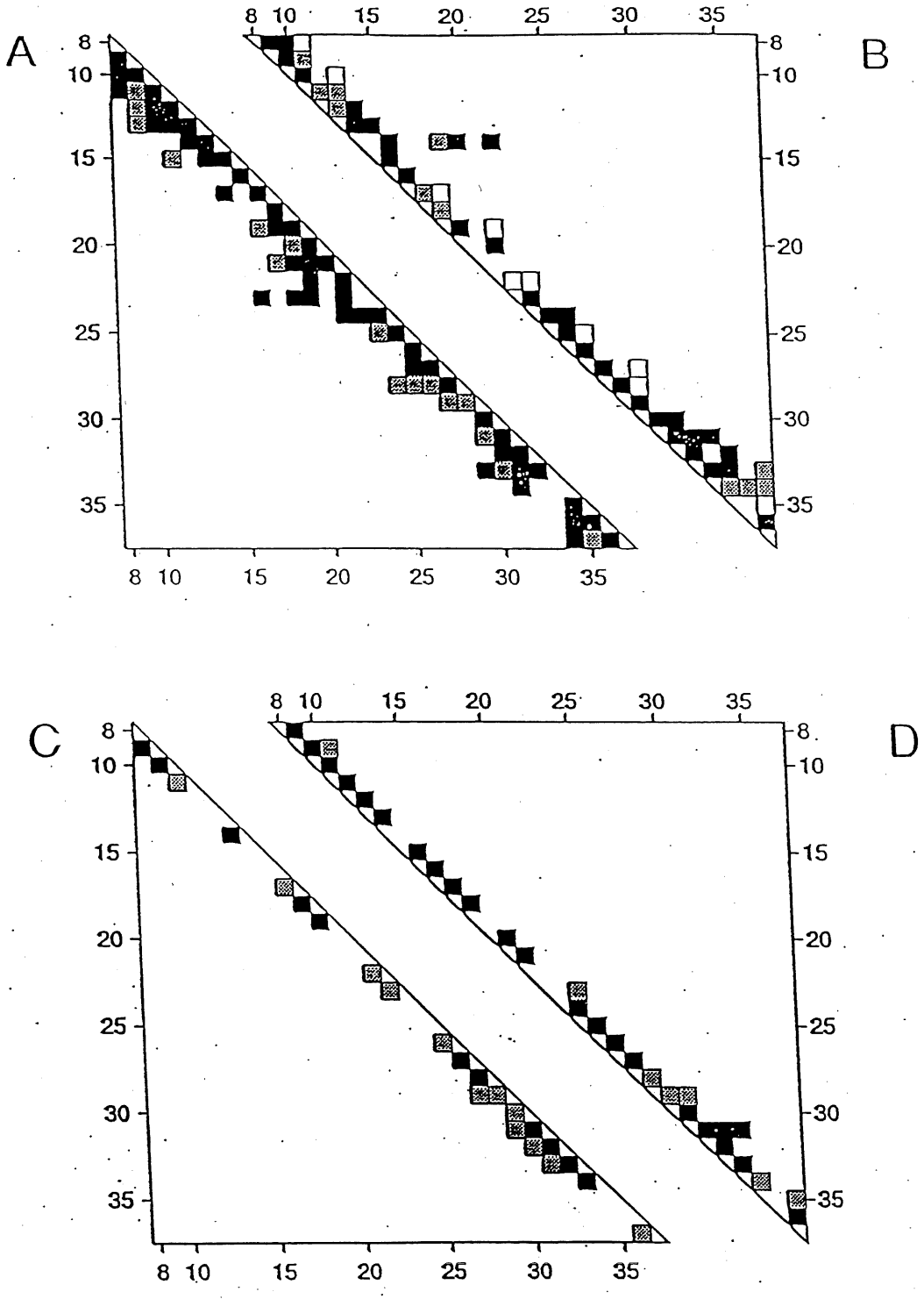


Fig. 1

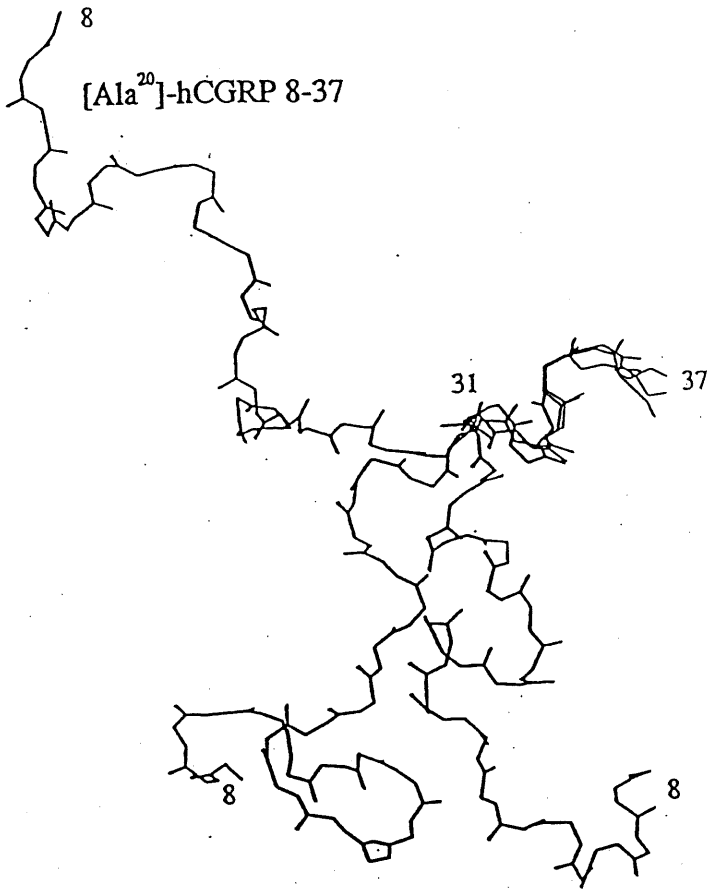


Fig. 2

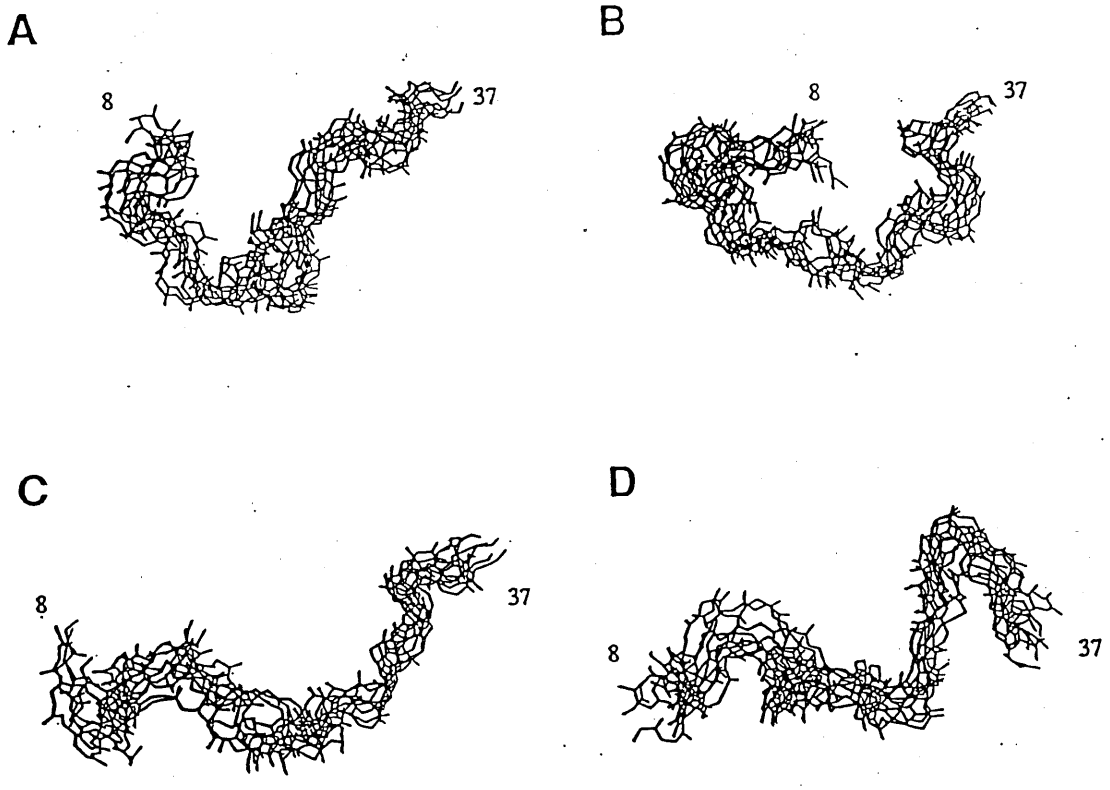


Fig. 3

CONCLUSION

L'objectif du travail de recherche proposé dans les chapitres II, III, IV et V était de traiter du problème de la relation structure-activité et de localiser, soit des différences structurales entre des peptides qui peuvent expliquer leur différence d'activité biologique, soit l'affinité structurale entre deux molécules qui peut expliquer leur association. D'après les résultats obtenus, nous pensons que l'objectif a été atteint.

Dans le chapitre I de ce travail, nous avons déterminé les structures tridimensionnelles du hCGRP 1-37 et de son antagoniste le fragment hCGRP 8-37 par RMN et modélisation moléculaire. Nous avons pu établir des différences structurales fondamentales qui peuvent expliquer le comportement biologique des deux peptides. Ainsi le segment en hélice (résidu 8-16) dans la structure du hCGRP 1-37 n'est plus présent dans la structure du hCGRP 8-37. L'explication avancée et la plus probable, est que l'hélice est stabilisée par des liaisons hydrogène avec le pont disulfure (résidu 1-7). Ces éléments de structures ont été confirmés par l'étude de l'effet de la température sur le déplacement chimique des protons NH.

Le chapitre V traite de la même famille de peptides. Ainsi les structures tridimensionnelles du hCGRP 8-37 et de trois de ses analogues Ala^{17}-, Ala^{20}- et Ala^{21}-hCGRP 8-37 ont été déterminées par RMN et modélisation moléculaire. Dans cette étude, nous avons pu démontrer que la structure du segment (Asn^{31}-Phe^{37}) se conserve dans les trois analogues hCGRP 8-37, Ala^{17}- et Ala^{20}-hCGRP 8-37, ce qui a permis de suggérer une explication pouvant être à l'origine de la diminution de l'activité antagoniste dans le cas de l'analogue Ala^{21}-hCGRP 8-37.

Les études antérieures effectuées sur la famille du peptide hCGRP ont indiqué que les résidus Arg^{11} à Val^{23} jouent un rôle conformationnel des plus importants dans l'interaction du hCGRP avec son récepteur. Une étude plus approfondie pourrait être effectuée en remplaçant spécifiquement chaque résidu de 11 à 23, en déterminant les structures tridimensionnelles des analogues résultants et en évaluant l'importance de chaque résidu dans le maintien des structures.

Dans le chapitre III, nous avons étudié les structures de la motiline 1-12 et de dix de ses analogues $(CH_2NH)^{1-2}$-,..., $(CH_2NH)^{10-11}$-motiline 1-12 à partir de l'effet de la température sur les déplacements chimiques des protons NH. Les structures tertiaires des trois analogues motiline 1-12, $(CH_2NH)^{1-2}$- et $(CH_2NH)^{4-5}$-motiline 1-12 ont été déterminées par RMN et modélisation moléculaire. À partir des résultats obtenus, nous pensons avoir identifié certains éléments qui peuvent être à l'origine, d'un côté de la réduction de l'activité dans le cas des analogues $(CH_2NH)^{1-2}$- et $(CH_2NH)^{2-3}$-motiline 1-12, et de l'autre côté de la perte totale de l'activité pour les autres analogues. Ces éléments sont les conformations adoptées par les segments Phe^1-Ile^4 et Thr^6-Arg^{12}. Des études plus poussées sont actuellement envisagées, afin de trouver des éléments de structures identiques entre la motiline et son analogue l'érythromycine A. En effet, des similitudes intéressantes ont été observées entre la structure de cette dernière et celle de la portion N-terminale de la motiline.

Dans le chapitre IV, nous avons étudié le complexe YSPTSPSY-ADN en utilisant deux approches en mécanique moléculaire. L'application de la méthode PEPSEA comme première approche, nous a permis d'analyser les populations de conformations accessibles au peptide YSPTSPSY, d'en sélectionner les plus stables et de les

comparer aux structures expérimentales. À l'aide de l'arrimage comme deuxième approche, nous avons étudié la formation du complexe YSPTSPSY-ADN et des arguments basés sur des données structurales et énergétiques ont été suggérés pour expliquer la formation et la stabilité du complexe. La comparaison de la conformation du complexe et des interactions intermoléculaires avec celle d'un complexe similaire résolu expérimentalement nous a convaincu de la validité et de l'intérêt des méthodes utilisées. Dans la continuité de ce travail, nous sommes actuellement en train de résoudre par RMN à deux dimensions les structures du peptide, du fragment d'ADN et du complexe, et ce, dans le but d'évaluer les résultats obtenus uniquement par modélisation moléculaire.

BIBLIOGRAPHIE

1. I.L. KARL. The peptides: analysis, synthesis, biology. V.4. Modern techniques of conformational, structural, and configurational analysis. Chap. 4, Édité par E. Cross et J. Meienhofer, Academic Press, New York (1981).
2. J. GUNNING et T. BLUNDELL. The peptides: analysis, synthesis, biology. V.4. Modern techniques of conformational, structural, and configurational analysis. Chap. 2, Édité par E. Cross et J. Meienhofer, Academic Press, New York (1981).
3. R.W. WOODY. The peptides: analysis, synthesis, biology. V.7. Conformation in biology and drug design. Chap. 2, Édité par V.J. Hruby, Academic Press, New York (1985).
4. P.W. SCHILLER. The peptides: analysis, synthesis, biology. V.7. Conformation in biology and drug design. Chap. 2, Édité par V.J. Hruby, Academic Press, New York (1985).
5. K. WÜTHRICH, G. WIDER, G. WAGNER et W. BRAUN. J. Mol. Biol. 169, 921 (1982).
6. K. WÜTHRICH, NMR of proteins and nucleic Acids. Wiley, New York, (1986).
7. G.M. Clore et A.M. Gronenborn. CRC Crit. Rev. Biochem. Mol. Biol. 24, 479 (1989).
8. G.P. GIPPERT, P.F. WRIGHT et D.A. Case. Biochem. Pharmacol. 74, 15 (1990).
9. K. WÜTHRICH. Science. 243, 45 (1989).
10. M.J.S. DEWAR et W.J. THIEL. J. Am. Chem. Soc. 99, 4899 (1977).
11. M.J.S. DEWAR. Further perspectives in organic chemistry. Elsevier. Amsterdam, p. 109 (1978).
12. P.K. WEINER et P.A. KOLLMAN. J. Comput. Chem. 2, 287 (1981).

13. B.R. BROOKS, R.E. BRUCOLLERI, B.D OLAFSON, D.J STATES, S. SWAMINATHAN et M. KARPLUS. J. Comput. Chem. 4, 187 (1983).
14. P. DAUBER-OSGUTHORPE, V.A.. ROBERTS, D.J. OSGUTHORPE, J. WOLF, M. GENEST et A.T. HAGLER, Proteins: structure, function and genetics. 4, 31 (1988).
15. G. NÉMETHY, M.S. POTTLE et H.A. SCHERAGA. J. Phys. Chem. 87, 1883 (1983).
16. F.A. MOMANY, L.M. CARRUTHERS, R.F. McGUIRE et H.A. SCHERAGA. J. Phys. Chem. 78, 1595 (1974).
17. B.R. GELIN et M. KARPLUS. J. Am. Chem. Soc. 97, 6696 (1975).
18. A.T. HAGLER, P.S. STERN, R. SHARON, J.M. BECKER et F. NADER. J. Am. Chem. Soc. 101, 6842 (1979).
19. R.A. DAMMKOELER, S.F, KARASEK., E.F.B SHANDS et G.R. MARSHALL. J. Comput. Aided Mol. Des. 3, 19 (1989).
20. M. VÀSQUEZ et H.A. SCHERAGA. Biopolymers. 24, 1437 (1985).
21. M. KARPLUS et G.A. PETSKO. Nature. 347, 631 (1990).
22. W.F. VAN GUNSTEREN. Curr. Opin. Struct. Biol. 3, 277 (1993).
23. B.R. BROOKS. Proceedings from supercomputer research in chemistry and chemical engineering. Édité par D. Truhler et K. Jensen. Am. Chem. Soc., Washington, DC (1987).
24. A.E. HOWARD et P.A. KOLLMAN. J. Med. Chem. 31, 1669 (1988).
25. N. METROPOLIS, A.W. ROSENBLUTH, A.H. TELLER et E. TELLER. J. Chem. Phys. 21,1087 (1953).
26. S. FRAGA et J.M.R. PARKER. Amino Acids. 7, 175 (1994).
27. S. KIRKPATRICK, C.D. GELAT, Jr., et M.P. VECCHI. Science. 220, 4598 (1983).
28. S.D. O'CONNOR., P.E. SMITH, F. AL-OBEIDI et B.M. PETTITT. J. Med. Chem. 35, 2870 (1992).

29. C.A. HASSANI. *La génération et l'analyse des populations de structures peptidiques en vue d'étudier les hypersurfaces d'énergie conformationnelle en mécanique classique.* Thèse de Doctorat, Université de Sherbrooke (1991).

30. T.F. HAVEL. Prog. Biophys. Mol. Biol. 56, 43 (1991).

31. T.F. HAVEL. Biopolymers. 29, 1565 (1990).

32. L.M. BLUMENTHAL. Theory and application of distance geometry, Chelsea, N.Y. (1970).

33. A.T. BRÜNGER et M. KARPLUS. Acc. Chem. Res. 24, 54 (1991).

34. T.L. BLUNDELL, B.L. SIBANDA, M.J.E. STERNBERG et J.M. THORNTON. Nature. 326, 347 (1987).

35. T.L. BLUNDELL, D. CARNEY, S. GARDNER, F. HAYES, B. HOWLIN, T. HUBBARD, J. OVERINGTON, D.A. SINGH, B.L. SIBANDA et M. SUTCLIFF. Eur. J. Biochem. 172, 513 (1988).